

Open Research Online

The Open University's repository of research publications and other research outputs

Modelling the transmission dynamics of RSV and the impact of routine vaccination

Thesis

How to cite:

Kinyanjui, Timothy Muiruri (2013). Modelling the transmission dynamics of RSV and the impact of routine vaccination. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2013 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:
<http://dx.doi.org/doi:10.21954/ou.ro.0000d594>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Modelling the Transmission Dynamics of RSV and the Impact of Routine Vaccination

by

Kinyanjui, Timothy Muiruri (BSc)

Affiliated Research Centre

KEMRI-Wellcome Trust Research Programme

Kilifi, Kenya

In collaboration with

University of Warwick, UK

March 2013

Thesis submitted for the Degree of Doctor of Philosophy

The Open University, UK

DATE OF SUBMISSION : 28 MARCH 2013

DATE OF ACCEPTANCE : 19 AUGUST 2013

BEST COPY

AVAILABLE

Variable print quality

Dedicated to

my loving wife (Beth Wambui)

Mum & late Dad

YOUR ACCEPTANCE

1 Student details

Your full name: Kinyanjui, Timothy Muiruri

Personal identifier (PI): A7330836

Affiliated Research Centre (ARC) (if applicable): KEMRI-Wellcome Trust
Research Programme, Kilifi

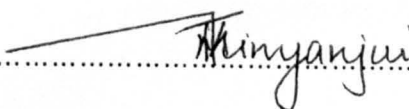
Department: Mathematics and statistics

Thesis title: Modelling the Transmission Dynamics of RSV and the Impact of
Routine Vaccination

2 Authorisation statement

I confirm that I am willing for my thesis to be made available to readers by The
Open University Library, and that it may be photocopied, subject to the
discretion of the Librarian

Signed:



Print name: Kinyanjui, Timothy Muiruri

Date: 06/09/2013 DD/MM/YY

3 British Library Authorisation (PhD and EdD candidates only)

The Open University has agreed that a copy of your thesis can be made
available on loan to the British Library Thesis Service on a voluntary basis. The
British Library may make the thesis available online. Please indicate your
preference below:

☒ I am willing for The Open University to loan the British Library a copy of my
thesis

OR

☐ I do not wish The Open University to loan the British Library a copy of my
thesis

Abstract

Introduction: Respiratory Syncytial Virus is the major viral cause of lower respiratory tract disease in young children worldwide, with the greatest burden of disease in infants aged 1-3 months. Consequently, vaccine development has centered on a vaccine to directly protect the infants in this age group. The fundamental problem is that these young infants are poor responders to candidate RSV vaccines. This thesis focuses on the use of mathematical models to explore the merits of vaccination.

Methods: Following development and analysis of a simple non-age-structured ODE model, we elaborate this to a Realistic Age Structured model (RAS) capturing the key epidemiological characteristics of RSV and incorporating age-specific vaccination options. The compartmental ODE model was calibrated using age-specific and time series hospitalization data from a rural coastal Kenyan population. The determination of Who Acquires Infection From Whom (WAIFW) matrix was done using social contact data from 1) a synthetic mixing matrix generated from primarily household occupancy data and 2) a diary study that we conducted in the Kilifi Health and Demographic Surveillance System (KHDSS). The vaccine was assumed to elicit partial immunity equivalent to wild type infection and its impact was measured by the ratio of hospitalized RSV cases after to before introduction of vaccination. Uncertainty and sensitivity analysis were undertaken using Latin Hypercube Sampling (LHS) and partial rank correlation respectively. Given the importance of households in the transmission of respiratory infections, an exploratory household model was developed to capture the transmission dynamics of RSV A and B in a population of households.

Results: From the analytical work of the simple ODE model, we have demonstrated that the model has the potential to exhibit a backward bifurcation curve within realistic parameter ranges. Both the diary and the synthetic mixing matrices had similar characteristics i.e. strong assortative mixing in individuals less than

30 years old and strong mixing between children less than 5 years and adults between 20 and 50 years old. When the two matrices were jointly linearly regressed, their elements were well correlated with an $R^2 \approx 0.6$. The RAS model was capable of capturing the age-specific disease and the temporal epidemic nature of RSV in the specified location. Introduction of routine universal vaccination at ages varying from the first month of life to the 10th year of life resulted in optimal long-term benefit at 7 months (for the diary contact model) and 5 months (for the synthetic contact model). The greatest benefit arose under the assumption of age-related mixing with the contact diary data with no great deal of effectiveness lost when the vaccine is delayed between 5 and 12 months of age from birth. Vaccination was also shown to change the temporal dynamics of RSV hospitalizations and also to increase the average age at primary infection. From the sensitivity analysis, we identified the duration of RSV specific maternal antibodies, duration of primary and tertiary infections as the most important parameters in explaining the imprecision observed in predicting both the age specific hospitalizations and the optimal month at vaccination. Results from the household model have demonstrated that the household epidemic profile may be different from the general population with strong interaction of the viruses in the household that do not necessarily reflect at the population level.

Conclusion: The synthetic matrix method would be a preferable alternative route in estimating mixing patterns in populations with the required socio-demographic data. Retrospectively, the synthetic mixing data can be used to reconstruct contact patterns in the past and therefore beneficial in assessing the effect of demographic transition in disease transmission. Universal infant vaccination has the potential to significantly reduce the burden of RSV associated disease, even with delayed vaccination between 5 and 12 months. This age class represents the group that is being targeted by vaccines that are currently under development. More accurate data measuring the duration of RSV specific maternal antibodies and the duration of infections are required to reduce the uncertainty in the model predictions.

Acknowledgements

One of the joys of completion is to look over the journey past and remember all the friends and family who have helped and supported me along this long but fulfilling journey. I would like to thank my supervisors Prof. James Nokes and Prof. Graham Medley for providing me with this great opportunity to explore the world of modelling infectious diseases. James for your patience and dedication in guiding me through the project as well as your attention to detail and critical analysis. Graham for your remarkable mathematical insights, prompt reply of numerous emails and always keeping my spirit high with your occasional humour. You both have helped me appreciate the art of constructive criticism. I would also like to thank Dr. Thomas House (University of Warwick) who has helped me a great deal with the Matlab code optimization, model fitting to hospitalization data, household model development and in general always sizzling with new ideas which are not only spectacular but helpful.

I would also like to thank Moses Kiti and the ever so dedicated fieldwork and data team that made the diary contact study a success. It indeed was not easy working on many a weekend consenting, issuing and collecting diaries from participants. Without you, it would have been impossible to complete this project and special thanks to all the study participants for their dedication and zeal of participation. The librarian at KEMRI-Wellcome Trust Kilifi (Alex Maina) - thank you so much for always assisting in sourcing many of the publications that I required and your invaluable tutelage on referencing and management of publications. I also acknowledge the help of John Ojal and Mark Otiende in extracting the demographic and the household occupancy data from the Kilifi Health and Demographic Surveillance System (KHDSS). Thanks to the KEMRI-Wellcome Trust training department for offering me the opportunity to carry out this project and for providing the necessary infrastructure. I would also like to thank my office mates Sande, Nyaigoti, Patrick, James, Patience, Moses and Joyce with whom I

have always discussed the preliminary results of this project and on many occasions as a result of the discussions, enhanced my understanding on the biology and epidemiology of RSV. Thank you very much for your time despite your busy schedules. Thanks to Dr. Lisa White (Mahidol-Oxford Research Unit, Bangkok) for the excellent discussions that resulted in wonderful outcomes.

Special thanks to my mum and late dad for always instilling in me the value of discipline, hard work and the fear of God. Thanks to my siblings (Njeri, Karuki and Cecilia) for their fascination with my subject of study and their constant questioning of its public importance. Kathleen and Simon, you made this project worthwhile. I would also like to appreciate my loving wife, Beth, who shares my passions, love and joy and who let me on many a night come home late during the course of writing up this thesis. Thank you for your love and unending dedication and I love you so much. And lastly, but by no means least, I would like to thank God for giving me this chance to glorify Him in this work and it is in Him that I live, move and have my being.

Contents

1	General introduction	1
1.1	Defining the question	1
1.2	Research objectives	6
1.3	Approach	7
1.4	Declaration of author’s role	8
1.5	Overview of the thesis	8
2	Biology and epidemiology of RSV	11
2.1	Historical overview	11
2.2	Antigenic variants of RSV	12
2.3	Infection, immunity and re-infection	12
2.3.1	Immunity to RSV	16
2.4	Characteristics of transmission	18
2.4.1	Seasonality	18
2.4.2	Transmission route	20
2.4.3	Risk factors for severe RSV disease	22
2.4.3.1	Extrinsic risk factors	22
2.4.3.2	Intrinsic risk factors	24

2.5	State of RSV vaccine development	27
2.5.1	Challenges to RSV vaccine development	35
3	Review of data on the epidemiology and transmission of RSV	38
3.1	Introduction	38
3.2	Objectives	40
3.3	Epidemiological parameters	40
3.3.1	Duration of RSV specific maternal antibodies	40
3.3.2	Rate of infection	46
3.3.3	Rate of RSV associated LRTI	55
3.3.4	Duration of infectiousness	65
3.3.5	Duration of immunity	78
3.4	Summary and discussion	81
4	Social contact patterns relevant for the spread of RSV	88
4.1	Introduction	88
4.2	Objectives	91
4.3	Methods	91
4.3.1	Prospective diary survey	91
4.3.1.1	Study design	91
4.3.1.2	Study site	91
4.3.1.3	Study population	93
4.3.1.4	Sample size calculation	94
4.3.1.5	Sampling procedure	97
4.3.1.6	Tool specification	98
4.3.1.7	Community engagement	99

4.3.1.8	Ethical considerations	99
4.3.1.9	Diary allocation and keeping procedures	100
4.3.1.10	Data capture, management and analysis	101
	Data capture and storage	101
	Data analysis	102
4.3.2	Synthetic mixing matrix	105
	Household contacts	106
	School contacts	107
	General/other contacts	107
4.4	Results	109
4.4.1	Diary survey results	109
	4.4.1.1 Age related mixing pattern	111
4.4.2	Synthetic mixing results	116
	4.4.2.1 Contact surface by administrative location	117
4.5	Discussion	123
5	A mathematical model of RSV transmission dynamics	129
5.1	Introduction	129
	5.1.1 Objectives	132
5.2	Model structure	132
5.3	Results	138
	5.3.1 Invasion threshold	138
	5.3.2 Local stability of the model	141
	5.3.3 Numerical simulations	146
	5.3.4 Sensitivity analysis	149

5.4	Summary and discussion	156
6	Modelling RSV transmission dynamics and the potential impact of vaccination	159
6.1	Introduction	159
6.1.1	Objectives	161
6.2	Methods	161
6.2.1	Model structure	161
6.2.1.1	Demographic sub-model	162
6.2.1.2	Epidemiological sub-model	163
6.2.1.3	Force of infection	166
6.2.2	Parameter estimates and model fitting	168
6.2.3	Vaccination	172
6.2.4	Uncertainty and Sensitivity analysis	174
6.2.5	Numerical techniques and presentation of the results	179
6.3	Results	180
6.3.1	Pre-vaccination and fitting results	180
6.3.1.1	Diary mixing model	180
6.3.1.2	Synthetic mixing model	184
6.3.2	Vaccination results	188
6.3.2.1	Vaccination of all susceptibles	188
6.3.2.2	Vaccination of primary susceptibles	200
6.3.3	Uncertainty and sensitivity analysis	202
6.3.3.1	U&S of the diary model	202
6.3.3.2	U&S of the synthetic model	207

6.4	Discussion	211
7	Modelling transmission dynamics in households	217
7.1	Introduction	217
7.2	Objectives	220
7.3	Methods	220
7.4	Results	227
7.5	Discussion	231
8	Final discussion	233
8.1	Introduction	233
8.2	Summary of the main findings	233
8.3	Limitations of the study and future research	241
8.4	Concluding remarks	245
	Bibliography	246
	Appendix A Diary sample	272
	Appendix B POLYMOD contacts for UK	273
	Appendix C Optimal sinks	275
	Appendix D Parameter table	277
	Appendix E Matlab code	279

List of Tables

2.1	Key target populations for an RSV vaccine	29
3.1	Estimates of the duration of RSV specific maternal antibodies . . .	43
3.2	Estimates of the rate of infection	51
3.3	Estimates of rate of RSV associated LRTI	59
3.4	Estimates of the duration of shedding	71
4.1	Reasons for non-consent	109
4.2	Number of recorded contacts per participant per day by different characteristics	112
5.1	Parameter estimates used in the numerical simulations	136
5.2	Parameter estimates for the baseline and the sensitivity analysis . .	154
6.1	Baseline parameter estimates used in the numerical simulations of the age-structured model	171
6.2	Baseline parameter estimates for the age-specific disease risk i.e. d_0, d_1 and d_2 following infection	172
6.3	Baseline parameter estimates for the age-specific risk of hospitaliza- tion following disease	173

- 6.4 Model parameters that have been included in the sensitivity analysis 178
- 6.5 Shows the parameters that were included in the fitting, their optimal values and the lower and upper 95% confidence limits for the diary contacts model 183
- 6.6 Shows the parameters that were included in the fitting, their optimal values and the lower and upper 95% confidence limits for the synthetic contacts model 187
- 7.1 Shows the baseline parameters used in the multistrain household model 226
- B.1 Contact matrix of all reported contacts in Great Britain consisting of the average number of contact persons recorded per day per survey participant in the POLYMOD study 274
- D.1 Shows the parameters and how they were used within the modelling work presented 278

List of Figures

2.1	Structure of RSV virion.	13
4.1	Study area within the KHDSS	92
4.2	Population pyramid of the KHDSS in 2011	94
4.3	Sensitivity measured as the variability observed in the force of infection when the daily contact rate between two classes e.g. infants and preschoolers, is varied between $\pm 20\%$	96
4.4	Distribution of contacts	110
4.5	Distribution of unique contacts	111
4.6	Mean number of people contacted per day per participant per age class	114
4.7	The daily rate of contact with different individuals accounting for reciprocity. The contact rate is represented as the mean number of contacts an individual in age class i has with respondents in age class j per day.	115
4.8	Household contact surface deduced from the Kilifi HDSS household occupancy data for the period from Sep 2010 to Jan 2011	116

4.9	Scatter plot of diary mixing matrix elements against the household mixing matrix elements.	118
4.10	Administrative locations in the KHDSS	120
4.11	Deduced patterns of household mixing within the different administrative locations within the KHDSS between Sept 2010 and Jan 2011	121
4.12	Distribution of household sizes by location for data collected between Sept 2010 and Jan 2011.	122
5.1	Schematic representation of the epidemiological model	133
5.2	Numerical solution of the model at disease free equilibrium and following the introduction infection	147
5.3	Bifurcation diagram showing the invasion threshold	149
5.4	Bifurcation sensitivity analysis	151
5.5	Shows backward bifurcation resulting from parameters shown in Table 5.2	155
6.1	Schematic representation of the epidemiological model	165
6.2	Shows the steps involved in generating the Latin hypercube sample	177
6.3	Model fit to RSV related hospitalizations for the diary contacts model fitted using diary contacts data	181
6.4	Bubble plot showing the distribution of hospital RSV related cases in each age class that are due to primary (red), secondary (yellow) and tertiary & subsequent infections (blue) for the diary model. . .	182
6.5	Model fit to RSV related hospitalizations for the synthetic model fitted using synthetic contacts data	185

6.6	Bubble plot showing the distribution of hospital RSV related cases in each age class that are due to primary (red), secondary (yellow) and tertiary & subsequent infections (blue) for the synthetic model.	186
6.7	Vaccination with a vaccine that protects all susceptibles i.e. S_0 , S_1 and S_2	189
6.8	Distribution of susceptibles and maternally protected at equilibrium in the presence of vaccination.	191
6.9	Distribution of infecteds, I_0 , I_1 and I_2 , at equilibrium in the presence of vaccination.	193
6.10	The profile of the force of infection at equilibrium both in the presence and absence of vaccination.	194
6.11	Proportion reduction in the number of hospitalizations at the optimal month for both the diary and the synthetic model accounting for both direct and indirect effects of vaccination	195
6.12	Shows the short-term temporal dynamics after introduction of routine vaccination	197
6.13	Shows the average age at primary infection (in years) as a function of both the vaccination coverage and the age at vaccination for the diary-based model	198
6.14	Shows the average age at primary infection (in years) as a function of both the vaccination coverage and the age at vaccination for the synthetic model	199
6.15	Vaccination with a vaccine that protects primary susceptibles i.e. S_0	200

6.16	Shows the prediction imprecision (diary-based model) in the number of age-specific (A) and time series hospitalizations (B) that is attributable to variation in the input parameters while (C) shows the PRCC per parameter for each of the age classes.	203
6.17	(A) shows the prediction imprecision in the optimal month at vaccination that is attributable to variation in the input parameters while (B) shows the PRCC for each of the parameters for the diary model.	206
6.18	Shows the prediction imprecision (for the synthetic model) in the number of age-specific (A) and time series hospitalizations (B) that is attributable to variation in the input parameters while (C) shows the PRCC per parameter for each of the age classes.	208
6.19	(A) shows the prediction imprecision in the optimal month at vaccination that is attributable to variation in the input parameters while (B) shows the PRCC for each of the parameters for the synthetic model.	210
7.1	Graphical representation showing the epidemiological compartments and the flow of individuals between them.	221
7.2	Coupling system used to represent the interaction between (blue arrow) and within (black arrow) households in the population. . . .	223
7.3	Epidemiological quantities as defined in the main text. A and C shows the transient and endemic equilibrium for the different infected classes at the population level while B and D shows the household dynamics at the peak and endemic prevalence respectively	227

- 7.4 The joint probability of observing a household with x RSV B and y RSV A infections at equilibrium and for both primary and secondary infections. 230
- A.1 Sample of the pictorial/text diary used during the survey 272
- C.1 Optimal sinks for the fitted parameters in the diary model 275
- C.2 Optimal sinks for the fitted parameters in the synthetic model . . . 276

List of Abbreviations

A_p	Average age at primary infection
R^2	Coefficient of determination
R_0	Basic Reproduction Number
R_e	Effective Reproduction Number
AA	Amino Acid
ARD	Acute Respiratory Disease
ARI	Acute Respiratory Infection
CAST	Community Advice for Specific Study Teams
CCA	Chimpanzee Coryza Agent
CI	Confidence Interval
CR	Case Ratio
cyo	Child years of observation
DFE	Disease Free Equilibrium

ELISA Enzyme-Linked Immunosorbent Assay

Eqn Equation

FGDs Focus Group Discussions

GA Gestational Age

IFAT Immunofluorescent Antibody Test

IgG Immunoglobulin G

INLA Integrated Nested Laplace Approximation

IR Incidence Rate

KDH Kilifi District Hospital

KEMRI Kenya Medical Research Institute

LHS Latin Hypercube Sampling

LRI Lower Respiratory Infection

LRTI Lower Respiratory Tract Infection

M-PCR Multiplexed Polymerase Chain Reaction

matAb Maternal antibody

MCMC Monte Carlo Markov Chains

MLE Maximum Likelihood Estimation

MSEIRS Maternally-protected Susceptible Exposed Infected Recovered Susceptible

NFS Nasopharyngeal Flocked Swab

NGM Next Generation Matrix

NPA Nasopharyngeal Aspirates

NW Nasopharyngeal Wash

ODE Ordinary Differential Equations

OR Odds Ratio

PATH Program for Appropriate Technology in Health

PDE Partial Differential Equations

PFP-2 Purified Fusion Protein RSV vaccine

PIV Parainfluenza virus

PRCC Partial Rank Correlation Coefficient

pyo Persons Year of Observation

qPCR Quantitative real-time reverse transcription-Polymerase Chain Reaction assay

RAS Realistic Age Structured

RR Relative Risk

RSV Respiratory Syncytial Virus

SDE Stochastic Differential Equation

SIR Susceptible Infected Recovered

SIRS Susceptible Infected Recovered Susceptible

SIS Susceptible Infected Susceptible

TR-FIA Time-resolved Fluoroimmunoassay

U&S Uncertainty and Sensitivity analysis

UK United Kingdom

URTI Upper Respiratory Tract Infection

USA United States of America

VLP Virus-like particles

WAIFW Who Acquires Infection From Whom

WCW Who Contacts Whom

WHO World Health Organization

Chapter 1

General introduction

1.1 Defining the question

Respiratory Syncytial Virus (RSV) was first recovered in 1955 from symptomatic chimpanzees and was known as the chimpanzee coryza agent (CCA) [135]. Soon after this discovery, Chanock et al [30, 31] isolated the virus in infants with severe lower respiratory tract infections (LRTI) and these findings revealed the magnitude of the role played by RSV in causing LRTI in infants and young children. Since its discovery, RSV has increasingly been recognized as one of the major viral pathogens causing extensive outbreaks in the infants and young children worldwide [145]. Work in the developing country setting has revealed that RSV predominates other respiratory viruses in the disease burden. Results from a study conducted in a number of developing countries show that viruses are the most common organisms recovered from children with LRTI, and RSV is the viral agent identified most frequently [185].

Most studies on RSV burden are based on hospital surveillance, which is use-

ful in measuring the relative prevalence of different agents in severe LRTI. Some hospital studies also use population denominators to estimate RSV incidence but they likely under-estimate the true incidence because of health care access issues and they obviously under-estimate infection since they record only diseased individuals. Community based studies provide a clearer picture of the true incidence of RSV and have identified RSV incidence of severe pneumonia to be at least four fold higher than that based on hospital surveillance [148]. Variable rates of infection due to RSV associated lower respiratory tract infections in developing countries has been reported and values range from 4/1000 cyo to 430/1000 cyo [43, 193, 150, 15, 171].

RSV has also been implicated in severe lung disease in adults especially the elderly and the immunocompromised [52]. However, the peak of severe disease is observed in young infants mostly between the age of 1-3 months. A fundamental characteristic of RSV infection is incomplete immunity following primary and repeated infections [68, 3, 93] which implies that an individual will remain partially susceptible to re-infection throughout life. Other key features of RSV include its seasonally recurrent epidemics, patterns of occurrence in relation to strain variation, and age distribution of disease in relation to primary infections and re-infections. This chapter describes further the epidemiological features of RSV that make this a fascinating virus to study, but also which lead to difficulties in identifying the mechanisms that underlie the observed dynamics and in predicting the potential impact of immunization. Hence the argued need to utilize a modelling approach to explore the merits of different vaccination strategies.

The pattern of RSV transmission has been shown to vary between different geographical regions but consistently displays a strong seasonal component [204,

149, 26, 32], the mechanism for which is not thoroughly understood. Meteorological conditions have been implicated [201] but insufficient evidence exists to support any such role. The availability of susceptible individuals to be infected, not only naive but also secondary susceptibles, is likely to play a fundamental role in RSV's epidemic dynamics [205]. Seasonal incidence is likely to be further complicated by differences in the social contact pattern relevant for the spread of an infection as it has previously been demonstrated for measles [12, 57]. A number of social contact studies have demonstrated that contact patterns are highly assortative with age [136, 133, 48, 110]. Given such mixing patterns between individuals, the force of infection (the per susceptible instantaneous rate of infection) could be subject to seasonal changes due to possible variation in the contact rates at different times of the year e.g. higher contact rates between children during school terms. However, this has not been quantitatively demonstrated for RSV and therefore remains an unknown.

Given that the most severe RSV disease tends to occur in young infants between 1-3 months, then the necessity would be for a vaccine that confers protection on this group. A live attenuated vaccine has been put forward as the preferred strategy [161, 5] for immunizing uninfected infants because it is considered unlikely to cause vaccine enhanced respiratory disease that had been observed in a formalin inactivated RSV vaccine recipients following subsequent natural RSV infection in children naive to RSV [33]. However, these early infants may fail to respond adequately to vaccination either because of immunologic immaturity or the suppression of the immune responses by maternally derived antibodies [38]. Early infants are particularly susceptible to reaction from live attenuated vaccines, thus necessitating high vaccine attenuation with subsequent doubtful immunogenicity

[114]. Alternative vaccine strategies should therefore be evaluated that target a different group of individuals in the population but that offer a level of indirect protection to the most susceptible infants. Given the complexity of the transmission dynamics of RSV and the hindrances to effective vaccination early in life, the problem of RSV control lends itself to mathematical treatment. Mathematical models have a long history of contributing to the understanding of the impact of mass vaccination programmes in childhood diseases [56, 147, 97, 47]. Mathematical models are powerful tools that allow us to apply targeted measures more effectively as well as help in understanding some of the important underlying transmission processes influencing the behaviour of a system [131, 164, 186].

A variety of vaccine strategy options can be envisaged and since it may not be plausible or even ethical to try them on human subjects during clinical trials, their appropriateness can be explored using mathematical modelling. Modelling has the benefit that one can explore several options in a risk free environment. In order to examine the potential effects of introducing an RSV vaccine in the population, a model is required that realistically describes the transmission dynamics of RSV concentrating on the characteristics that are important in RSV epidemiology. There have been previous attempts to model the transmission of RSV. A comparison of a standard SIRS (Susceptible Infected Recovered Susceptible) model with a more realistic model of RSV transmission in which individuals acquire immunity gradually after repeated exposure has been undertaken [201]. The two models described the temporal dynamics equally well with the transmission parameter between the two models differing by a factor of 4. White et al [205] proposed a nested model that captured four possible host responses namely: partial susceptibility, altered duration of infection, reduced infectiousness and temporal immunity

to infection. The best fitting model was one where an individual remained partially immune during their lifetime. Neither of the two models evaluated the effect of introducing vaccination since they were not appropriately structured i.e. they did not include age-related heterogeneities that are essential for evaluating the benefit arising from vaccination. There has been only a single modelling study looking at the effect of introducing RSV vaccine into a population. Acedo et al [1] developed an SIRS model with two age classes, infants aged 0-1 years and the second age class was composed of all other individuals, and applied vaccination at birth. However, this model has a number of limitations: it assumes 1) no maternal antibody protection at birth; 2) that mixing between the two age classes is homogeneous; 3) that primary and secondary infections have got the same recovery rates and level of infectivity; 4) that vaccination of infants at birth is 100% effective; and lastly, having just two age classes in the population is just a gross over simplifying assumption. These simplifying assumptions can lead to under or over estimation of the outcome of vaccination. All of these previous RSV models have compartmentalized the population exclusively in terms of the infection status and history excluding age heterogeneities. They also assume homogeneous mixing within the population ignoring differences in transmission potential despite evidence that most infections happen through a limited set of contacts [115, 136].

In the work presented in this thesis, I seek to address the limitations of the previous RSV models by developing a Realistic Age Structured (RAS) RSV model accounting for heterogeneity in the transmission by having age-specific rates of infection, acquisition of disease and hospitalization. Further, the model will account for reduced susceptibility and reduction in infectiousness of individuals according the number of previous infections. The model construction and parametrization

draws heavily from data sources within a Kenyan coastal population that has been extensively studied [182] for RSV for over a decade. We will then use the model developed to explore the effect of introducing routine RSV vaccination in the population. Additionally, we have introduced a simple household multi-strain RSV model to describe the transmission of RSV both within the household and the general population. Transmission of RSV within the household is an important mechanism for spread due to the greater strength of contacts between individuals sharing living arrangements compared to contacts outside of the household [122].

1.2 Research objectives

The overall objective of this study was to develop a mathematical model to describe the transmission dynamics of RSV within a developing country population, with which to explore the potential impact of different vaccination strategies on severe RSV disease in infants and young children.

Specific objectives

1. Develop basic understanding of the transmission dynamics of RSV through the use of both simple and realistic age-structured deterministic ODE (Ordinary Differential Equation) models.
2. Estimate, using two different approaches, the contact rate for a rural coastal population in Kenya by which to define the who acquires infection from whom (WAIFW) matrix for a realistic age-structured model for RSV.
3. Investigate the impact of introducing immunization into a developing country

population on the RSV associated number of hospitalizations.

4. Demonstrate using the household model that more epidemiologically relevant parameters are potentially identifiable compared to the mean field models as well as form a template on to which further biological complexity can be added.

1.3 Approach

A mathematical model is constructed to which an increasing level of biological and demographical complexity is added. The approach that we have taken of gradually building up the model complexity facilitates our understanding of the individual elements and to know to what extent they contribute to the observed patterns of transmission. We then introduced routine vaccination into the model explicitly exploring the effects of altering the age at which the vaccine is administered throughout infancy and early childhood. Routine vaccination was implemented as the continuous vaccination of a proportion of individuals at all time points passing a specified age gate. Sensitivity analysis of the model predictions to uncertainty in the input parameters is investigated using Latin Hypercube Sampling (LHS) and the importance of each of the parameters is explained using Partial Rank Correlation Coefficient (PRCC). This kind of incremental building of the model and the sensitivity and uncertainty analysis not only allows for the description of the transmission dynamics and the effect of introducing vaccination in the population to be assessed but also gives information about which model parameters require more precise data estimates. The model developed was based on an extensive review of data on the biology and epidemiology of RSV. Model parameters that could not

be identified from the published literature were estimated by fitting the model to RSV temporal and age-specific hospitalization data from the Kilifi District Hospital (KDH) using the rigorous maximum likelihood method. For the household model, we have considered a population grouped into households where people retain their random interactions within the population but suffer an additional rate of infection for every infectious person in the household. The model consists of a fairly large but simply described set of ODEs where the stochastic nature of the transmission is captured by modelling all the possible household configurations.

1.4 Declaration of author's role

The studies presented in this thesis were designed by the author together with supervisors, Prof. James Nokes and Prof. Graham Medley. The author of this thesis was responsible for the protocol development, data analysis, the development of the mathematical model and the running of the simulations. For the contact study, the author was the lead investigator for the protocol development and the subsequent analysis of the data arising from the work. The diary study was a collaborative effort with the supervisors and Moses Kiti who was responsible for the day-to-day field activities including task allocation to the field workers.

1.5 Overview of the thesis

Following this introductory chapter, I have presented two review chapters. Chapter 2 gives an introduction to the biology, the characteristics of transmission, and the status of RSV vaccines under development. Chapter 3 presents a review of

some of the most important parameters for the characterizing the epidemiology of RSV and discusses the limitations of the data. Chapter 4 gives an account of a study designed to define age-specific contact rates, used in later modelling, through the collection of daily diary recordings from a sample of individuals from the Kilifi Health and Demographic Surveillance System (KHDSS). The chapter also introduces a novel method that we used to generate a synthetic mixing matrix which is based principally on household occupancy data from the KHDSS. This enables comparisons in a later chapter to be made in RSV transmission and vaccine control based on two different formulations of age-specific contact rates. Chapter 5 introduces a simple RSV mathematical model by which to investigate some basic characteristics of RSV transmission dynamics. Useful statistics such as the basic reproduction number (R_0) and the invasion threshold have been presented. Multiple supercritical endemic equilibrium points have been shown to exist using backward bifurcation curves and may potentially influence the outcome of vaccination. This model has been extended in Chapter 6 to include age heterogeneities in transmission enabling an investigation of the potential impact of RSV infection and disease from age-specific universal vaccination. Sensitivity and uncertainty analysis, of this realistic age structured model, has also been presented by which to assess the robustness of the model predictions and identify the most important epidemiological parameters contributing to the uncertainty observed in the model outcomes. Chapter 7 contains the household model which we have used to demonstrate that the household models can potentially be used to determine more epidemiological parameters by distinguishing infection pattern within the household what is possibly indistinguishable at the population level. The final chapter presents a brief summary and discussion of the main findings from the

work in the thesis and discusses the potential directions for future research.

Chapter 2

Biology and epidemiology of RSV

2.1 Historical overview

Morris et al [135] observed sneezing, coughing and nasal discharge from a group of young chimpanzees in 1955. When they cultured nasal specimen from the symptomatic chimpanzees, they recovered an unrecognized virus which they named chimpanzee coryza agent (CCA). They further reported that human beings, particularly adolescents and young adults, have antibodies in their sera directed against the coryza agent suggesting that they have experienced infection with the agent or one that is closely related. Soon after this discovery, Chanock et al [31, 30] isolated CCA like virus in infants with severe lower respiratory tract infections. These agents were shown to be serologically related and of note was their ability to induce syncytia and multinucleated cells in Chang cells and therefore respiratory syncytial virus was suggested as a more appropriate name. Following its discovery, RSV has been identified as one of the major viral pathogens causing extensive outbreaks in the very young [145] and vulnerable adults [53, 197].

2.2 Antigenic variants of RSV

RSV is a member of the genus *Pneumovirus* in the family *Paramyxoviridae*. It has a negative sense, non-segmented single-stranded RNA genome comprising of 10 genes and encoding 11 proteins. These include genes for the attachment glycoprotein (G), the fusion protein (F), the small hydrophobic protein (SH), three matrix proteins (M1, M2-1 and M2-2), three nucleocapsid proteins (N, P and L) and two non-structural proteins (SH1 and SH2) [192, 23]. The two non-structural proteins are expressed only during cell infection and are not packaged into the virion. See Figure 2.1 for the structure of RSV. Based on serological reactivity with monoclonal antibodies, two antigenic groups, A and B have been identified [138, 6] and the data does suggest that the two sub-types may have evolved separately for a considerable time period. The main difference in the two subtypes was observed in the G protein which has also been reported to be the most variable protein [108] as demonstrated by glycoprotein specific assays of antibody responses induced by RSV infection of respiratory tract of cotton rats. It has also been shown that RSV A and B have 53% amino acid (AA) homology between prototype strains of the G protein [109] while the F protein has got 89% AA homology [107] between prototype strains. In addition to group variability, it has been noted that numerous strains and designated genotypes exist within both groups [3, 4, 24].

2.3 Infection, immunity and re-infection

RSV has a very high rate of transmission in the first few years of life. Studies show in excess of 50% of newborns are infected in their first year (first epidemic)

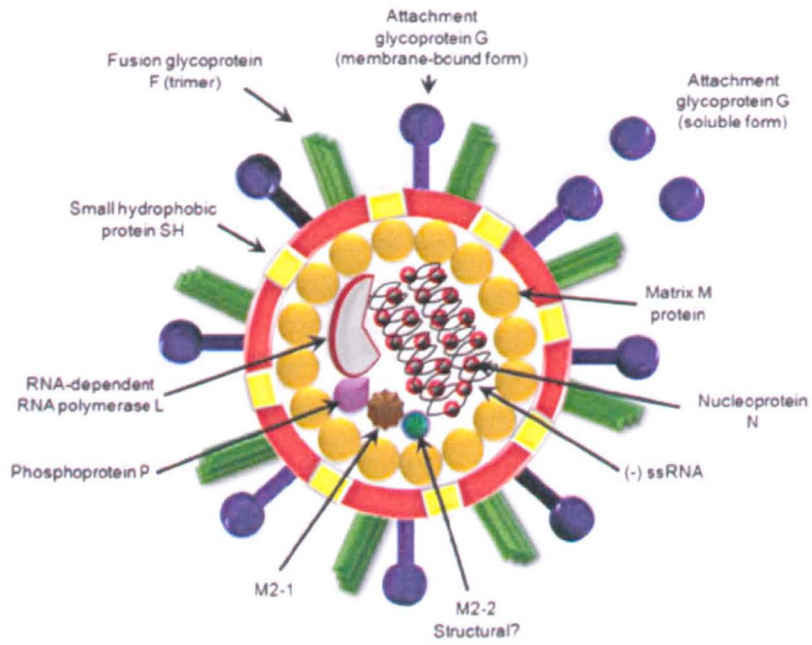


Figure 2.1: Figure showing the structure of RSV. The fusion protein (F), the attachment glycoprotein (G) and the small hydrophobic protein (SH) are located on the surface of the virion. The remaining structural proteins reside inside the envelope. Figure adapted from Gonzalez et al [70]

and the vast majority of the remainder in year two (second epidemic) [68, 93]. For example, in a classic study from the USA in which a cohort of infants were intensively monitored over the first five years of life showed that 68% (95% CI: 59.6 - 75.7) of the children contracted primary infection in their first year of life while 98% (95% CI: 87.1 - 99.8) of the remaining contracted primary infection during the second year of life [68]. Worldwide, it is estimated that 33.8 million (95% CI: 19.3 - 46.2) new episodes of RSV associated lower respiratory tract infection occurred in children ≤ 5 years between January and December 2005 with at least 3.4 million (95% CI: 2.8 - 4.3) episodes requiring hospitalization [145] indicating that the global burden of RSV is huge. A more detailed review of the rates of RSV infection can be found in Chapter 3.

RSV infection is usually symptomatic in children less than 5 years of age [150], almost invariably so in infancy [87] and may result in disease requiring hospitalization in the very young children, mostly those in their first year of life [150, 68]. There exists evidence from a number of studies that RSV in older children and adults is mostly associated with mild infection [87] but in some cases, there are reports of moderately severe disease [88, 52]. Hall et al [88] prospectively evaluated healthy adults aged between 18 and 60 years for respiratory virus infections between 1975 and 1995. Of the total number of individuals, 211 individuals (7%) acquired RSV infection with 84% of the subjects being symptomatic. Upper respiratory tract infection was observed in 74% of the individuals while 26% had lower respiratory tract symptoms. Overall, 40% of the subjects were febrile. These findings have been corroborated by a human experimental study that observed symptomatic infections in adults inoculated with RSV [89].

The risk of developing lower respiratory tract infection following primary in-

fection has been shown to decline in community studies. This evidence for decline in the risk of disease has been demonstrated by Glezen et al [68] where it was observed that this risk decreased from 22.4/100 cyo in children less than one year of age to 7.7/100 cyo in children aged 37-48 months. The risk of disease was zero in children over 49 months. Henderson et al [93] has demonstrated that age is not the only important factor but history of infection as well influenced the disease outcome. The authors noted that the attack rate of the first infection was 98% while that of the second and third infections respectively declined to 75% and 65% respectively. Half of the children who experienced second infections had less severe disease compared to those with primary infection while 14% had more severe manifestations. When third were compared with second infections, 38% were less severe while only 4% had more serious disease. A more recent study quantifying the effect of the history of infection on the risk of RSV disease i.e. severe lower respiratory tract infection, reported that the risk of disease for reinfection was lower relative to primary infection i.e. a reduction in the risk from 14.1% to 4.8% [155]. It is difficult to separate and identify the independent effect of age and history of infection on the expression of disease since reinfection and age are highly correlated and therefore it is likely that the two factors interact to modify the disease outcome.

More recently, pathogenicity of RSV has been demonstrated in elderly adults [90, 52, 88], institutionalized individuals [53] and those with compromised immune function [197, 52]. The importance of RSV infection in the elderly is increasingly being recognized. RSV associated pneumonia in individuals 65 years of age or more has been estimated to cause 14,000 - 62,000 hospitalizations each year in the United States of America. At a rate of 40 - 180 hospitalizations per 100,000 in

individuals older than 65 years of age, hospitalization is estimated to cost \$150 - \$680 million each year in the USA [90].

2.3.1 Immunity to RSV

It has been established that RSV can reinfect throughout life. In fact, in the family study conducted by Hall et al [87], significant attack rates were seen in all age groups where the attack rate in adults was reported to be 16.8% compared to 29.4% in infants. The study by Henderson et al [93] reports the attack rate for second infections in the second year of life to be 74.5% while that of third infections to be 65.4%. Following this, the study by Glezen et al [68] reported an attack rate of second infections to be 75.9/100 cyo (95% CI: 55.6 - 100.6). These findings taken together point out to the fact that the immune response elicited by a natural infection is incomplete and hence the common occurrence of re-infections throughout life.

The immune response to RSV is both humoral [36, 78] and cell mediated [159]. A number of studies [149, 68, 85, 21] have observed that disease occurs during early infancy when maternal antibodies are universally present although some studies have however shown that high levels of RSV specific maternal antibodies can potentially offer a protective effect against RSV infection [172, 154]. The protection provided however, appears to be partial and of limited duration. The high incidence of infection in the very young does seem to support the hypothesis that maternal antibodies do not offer complete protection. Parrott et al [160] have in fact argued that one of the factors that seem to contribute to an impaired immune response is immunologic suppression produced by maternally derived antibodies

in serum. Another study conducted by Ogilvie et al [154] has shown that the severity of illness from RSV infection in the first year of life is modified by high levels of maternal immunity. The authors do suggest that three possible mechanisms of protection may have been at play 1) the mother herself is protected from re-infection (a potential major source of infection for the infant), 2) maternal immunity may be transferred through colostrum or breast milk and 3) humoral antibody transferred across the placenta. More recently, it has been shown that prophylactic administration of RSV immunoglobulin given to high risk infants and young children is effective in reducing severe disease [76].

Similar to the case of maternal antibodies, the presence of acquired antibodies in individuals who have had a previous infection does not seem protective against infection although the risk of re-infection has been associated with the number of previous re-infections [87] and the level of pre-existing antibody [162]. Lack of antibody mediated protection could be due to the fact that RSV-blocking antibodies are usually elicited at very low frequencies after exposure sometimes requiring years of repetitive exposure to the virus for their acquisition [36]. An interesting observation by Cane et al [25] is that serum antibody responses are closely linked to the genotype of the infecting virus. This would make it possible to determine the genotype of the infecting strain even in individuals for whom no virus isolation is available but only on primary exposure when the individual is naive. Upon re-exposure with a different variant, there is a broadening of antibody response and boosting of antibodies to previous infecting strains [183]. Sande et al [178] have also demonstrated that serum infant and young child neutralizing response to RSV is significantly group specific (but not genotype specific) with this pattern of homologous vs heterologous reactivity similar irrespective of whether the test

viruses were contemporary or historical.

Given that repeat infections are observed even in the presence of effective neutralizing antibodies, it is likely that RSV infection is cleared by cell-mediated immune response which acts by directly destroying infected cells or by indirectly limiting inflammation in the lungs [159]. Healthy adults have been shown to display T-cell activation levels that are sufficient to mediate viral clearance from the lungs which promotes low levels of inflammation and hence reduced tissue pathology [159]. From the studies discussed in this section, it seems that one of the major factors driving re-infection even in adulthood is the lack of a complete immune response.

2.4 Characteristics of transmission

2.4.1 Seasonality

The transmission of RSV is characterized by a pronounced seasonal pattern in both the tropical [32, 168] and the temperate regions [204, 190]. Epidemics usually occur annually with the exceptions of some Scandinavian countries e.g. Finland which experiences a minor peak in April followed by a major peak in December [204] on a biennial basis i.e. double epidemic every two years. In general, seasonality is characterized by a duration of high transmission followed by a period of fade out with the exception of some tropical countries e.g. Singapore and Hawaii, where RSV has been observed to be present all year round [32, 168]. The reason for these periodic outbreaks are not clear although a number of factors have been postulated [71]. In a review by Weber et al [204], RSV outbreaks were more frequently

associated with rainy weather than the colder season with RSV peaking one or two months after the onset of the rainy season. In desert climates e.g. Kuwait [98] and Saudi Arabia [106], it has been reported that more cases are seen in the colder months. However, countries with perennial high rainfall exhibit a less clear cut seasonality pattern for example in Singapore and Hawaii where RSV is reported through out the year [32, 168]. In these regions, RSV cases are observed to be more in one half of the year compared to the other suggesting that perhaps the cyclical pattern is at least partly driven by biological interactions between virus and host. In the temperate climates [127], RSV peaks mainly during the winter months and this scheme appears to be independent of the rainfall pattern as winter has high rainfall in places such as Santiago in Chile [11] yet low rainfall in places such as Johannesburg in South Africa [111]. From the findings reported in the studies above, it seems that neither temperature nor rainfall is the major determinant of the timing of the RSV epidemics observed. However, climatic and geographical factors do seem to play a role but not exclusively, indicating that there may exist other factors, e.g host or social behavioral that influence the timing of the epidemics. Social behavioral factors such as school vacations and indoor crowding have been hypothesized to play a role [156]. Studies carried out have suggested that infants most likely acquire infection from school going children within the same household [156, 134, 87] with the school year being proposed as a possible mechanism driving the seasonality [200]. The seasonal pattern is also very likely influenced by the short lived immunity to RSV [3] which may translate into a build up of susceptibles who are available for re-infection with the addition of new births creating individuals who have lost maternal antibody protection. It is also worth noting that from a methodological point of view, most of the studies

reported are passive and therefore rely on the observability of subjects seeking medical attention from a health facility. It is possible that even in the absence of cases in the health facilities, transmission within the population is on going but without severe cases being observed.

2.4.2 Transmission route

It has been observed that RSV is highly transmissible which is evidenced by the rapidity with which it occurs following birth with 68% (95% CI: 59.5 - 75.7) or more new borns becoming infected in their first year of life [68]. It has also been shown that re-infections are common in older children and in adults [88]. It therefore seems intuitive that RSV transmission from one individual to another is highly effective. A number of studies have been carried out that have attempted to generate evidence on the route of transmission of RSV.

Hall et al [83] did a study to test whether nosocomial spread of RSV could occur through contact with environmental surfaces contaminated by RSV infected nasal secretions. RSV could be recovered from countertops and gloves for the longest period, an average of 7 and 5 hours respectively. On cloth, the infectious virus was recovered for an average of 2 hours while on skin and paper tissue, survival was diminished to an average of 30 minutes. RSV was recovered from hands contacting surfaces (cloth and tissues paper) contaminated by fresh secretions of infected infants for an average of 3-10 minutes while from countertops the average duration was 20-25 minutes. These findings indicate that RSV may survive sufficiently long in the environment to allow transfer of infectious virus to hands in contact with contaminated surfaces suggesting that such a route of transmission is plausible.

Hall et al [86] carried out another study to determine the possible spread of RSV to young adults working in a pediatric ward. Volunteers were divided into three groups with each of the group exposed to at least one route of transmission. The routes of transmission are a) by large particles or droplets, b) by self-inoculation after touching contaminated surfaces and c) by small particle aerosol. Infection was reported in the *cuddlers* (volunteers who cared for the babies) and *touchers* (volunteers who were only allowed to touch surfaces contaminated with the baby's nasal secretions) but not in the *sitters* (volunteers exposed to the infant by sitting a distance of more than 6 feet from an infant's bed) group. Additionally, all the *cuddlers* infected developed upper respiratory tract symptoms. This finding suggests that the spread of RSV may occur by close contact with direct inoculation of large droplets or by self-inoculation after touching contaminated surfaces. The inoculum obtained through direct inoculation of large particles may be greater than that received by contact with contaminated surfaces since the amount and survival of infectious virus is dependent upon the type of surface [86].

The feasibility of transmission by fomites has been demonstrated for rhinoviruses. Hendley et al [94] showed that rhinovirus would survive on environmental surfaces and skin for ≥ 1 hour and that subjects could infect themselves by touching their nasal or conjunctiva mucosa. Since transmission by fomites has also been shown to be a possible transmission route for RSV, then the importance of this route of transmission in the natural setting will depend on a) the titre of the virus in the secretions b) the minimal infectious dose required to allow for virus replication in the human host and c) the relative sensitivity of the route of infection. However, the inoculum received through direct inoculation of large mucosal particles may be greater than received by contact with contaminated particles since the virus is

extremely labile when exposed to the environment and the type of surface. Direct inoculation may therefore be more important as a route of transmission than transmission by fomites.

2.4.3 Risk factors for severe RSV disease

Pneumonia is a major cause of morbidity and mortality in the developing world and respiratory viruses make a significant contribution to this disease burden [185]. Among the viruses, RSV is one of the main contributors to community acquired pneumonia [149, 145]. However, RSV infection leads to generally mild illness of the upper respiratory tract but in a few patients, it will progress to severe disease requiring hospitalization. In this section, the main risk factors that influence the incidence of RSV-LRTI are considered and categorized into both intrinsic and extrinsic risk factors.

2.4.3.1 Extrinsic risk factors

A number of studies have identified crowding as a risk factor to RSV-LRTI. In the Tucson cohort study, the multivariate analysis showed a significant effect of the number of persons sharing a bedroom (RR (Relative Risk) , 4.0: $p < 0.002$) on the development of RSV-LRI [99]. In a birth cohort intensively monitored for RSV over 3 RSV epidemics identified that crowding (measured by the number of children in the home) and the number of children under six years in the home were found to correlate with increased risk of RSV-LRTI [157]. Another study investigating the risk factors for severe RSV in The Gambia identified that increased risk of RSV-LRTI was associated with greater numbers of children in the age group $3 - \leq 5$

years living in the same compound (OR=9.1 : 95% CI: 3.7-28 for ≥ 2 children in the age group $3- \leq 5$ years) [202]. In a case control study examining the risk factors of severe RSV in Alaskan children reported that having 4 or more children < 12 years in a household (OR=2.13: $p=0.01$) and a household crowding index of ≥ 2 (OR=1.72: $p=0.024$) were significant independent risk factors to severe RSV disease [22]. The household crowding index is defined as the total persons in a household divided by the number of rooms excluding bathrooms, hallways, closets. The actual mechanism through which crowding increases the risk is not well understood but it is hypothesized crowding is related to higher viral load either through interpersonal transmission or through a greater possibility of exposure to RSV in the first year of life through multiple persons in the household [188].

It has been established that respiratory illnesses are increased in infants and children attending day care groups outside of the home [191, 103]. A study by Strangert et al [191] reported that 1 year olds in day care centers had more febrile illnesses per child and more days of rhinitis alone than matched children in home care ($p<0.05$). In a case control study done in Atlanta [7] that examined the role of risk factors in hospitalized children (the most common causative agent associated with illness was RSV) suggested that regular attendance in a day care center with more than six children was an important risk factor for severe disease requiring hospitalization (OR, 2.08: $p<0.05$). A study by Liese et al [120] did not find day care attendance of the child as a significant factor but day care attendance of a sibling was identified as a significant independent predictor of RSV illness requiring hospitalization (OR, 3.9 95% CI, 1.9-8.3, $p<0.001$).

Passive smoke exposure has also been identified as a risk factor to the development of severe disease. In a case-control study carried out in Istanbul, Turkey

[77], the authors enrolled 28 infants with bronchiolitis and 30 in the control group. In this study, they demonstrated that serum cotinine levels of 10.8 mg/mL in the bronchiolitis cases were significantly elevated compared to controls with 3.9 mg/mL ($p < 0.0001$). There were also significantly more parents smoking in the households of children hospitalized with bronchiolitis compared to control subjects (82% vs 70% $p < 0.05$). In this study, no other risk factors were examined and therefore a multi-variate analysis was not performed. In a Spanish longitudinal observational cohort study carried out for two years spanning 1998 to 2000 that enrolled all premature infants ≤ 32 weeks GA (gestational age), observed in both the multivariate and univariate analysis that tobacco smoke exposure was a significant risk factor for development of severe RSV disease requiring hospitalization [27]. A study carried out in rural coastal Kenya to identify the risk factors associated with increased risk of progression to RSV associated pneumonia identified that exposure to tobacco smoke was a significant risk factor [157]. However, there exist other studies that have not identified exposure to passive smoke a significant risk factor [120, 99]. The reason for the inconsistent evidence is probably because the amount of smoke exposure is critical. However, the correlations reported have been made by measuring actual exposure rather than the extent of smoke exposure. In future studies, there may be merit in measure the amount of exposure rather than just exposure alone.

2.4.3.2 Intrinsic risk factors

Birth during the first half of the RSV season is a risk factor for the development of RSV lower respiratory tract infection. In a study consisting of infants enrolled into Tucson children respiratory study, the investigators reported that the incidence of

RSV-LRTI was higher for children born from July to December (16.2%) compared to those born from January to June [99]. Since most disease occurs in infants early in life, it is not surprising that birth during the first half of the epidemic season is associated with increased risk of RSV LRTI. Children born early in the epidemic are more likely to get infected in the epidemic, since there is a longer window for RSV infection, and hence get early infection when they are at the highest risk of developing severe disease compared to children born later in the epidemic (less time to get infected) who may not become infected until the second epidemic and hence they would be older. It is also possible that babies born early in the RSV epidemic, they have got low levels of RSV specific maternal antibodies since their mothers have not been exposed to the infection during pregnancy. Babies born in the later half of the RSV season may have had boosted levels of maternal antibodies since their mother may have had a re-infection. Boosted levels of maternal RSV specific maternal antibodies have got a potential of reducing the risk of infection or disease when one gets infected.

It is clear from a number of studies [68, 148, 145] that severe disease is most likely in early infancy following infection. A prospective surveillance study of severe and very severe pneumonia in children aged ≤ 5 years admitted from 2002 to 2007 in a rural district hospital in coastal Kenya observed that approximately 55% of all severe and very severe cases were reported in children ≤ 6 months of age [148]. In these infants, the absence of protective maternal antibodies [152, 153, 67], narrow airways [143] and an immature immune system [39] may play a significant role in predisposing them to more severe RSV disease.

The role of breast feeding has been established in reducing the frequency and the duration of LRTI [123]. In the Tucson study [99], it is reported that the OR

for having a RSV-LRTI in those infants of mothers with a low educational level who were not breast fed compared to those who were breast fed was 6.8 (95% CI: 0.8-56.0). Although a majority of case control studies in the developed world have reported that breast feeding protects against RSV disease [99, 123], others studies exist that have not shown this protection [7]. The risk factor study in The Gambia by Weber et al [202] has also not shown the protective role of breast feeding in the development of severe RSV. One of the possible reasons for the inconsistencies in the studies reported is in the definition of breastfeeding. The studies showing no effect did not define exclusivity in breastfeeding. The criteria for determining the effects of breast feeding should be strictly applied e.g. definition of exclusive breast feeding and controlling of other risk factors in the multivariate analysis, otherwise the role of protective effect might seem diminished. The possible biological reason that would result in breast milk being protective is the presence of lactoferrin and anti-RSV IgA in colostrum [92, 91, 38]. It has also been proposed that breast milk promotes lung maturation through prolactin [91].

Three possible avenues have been proposed for the control of RSV: prevention through vaccination, case management through supportive treatment and mitigation of risk factors [202]. In an ideal situation, it would be desirable to implement a combination of the three methods. Which of the methods or combination of methods to implement in resource poor settings is subject to a number of factors. Vaccination is likely to be the most effective method but so far there are no licenced RSV vaccines. Supportive treatment e.g. oxygen therapy, cohorting to reduce the nosocomial spread of RSV or prophylactic immunization of high risk infants using palivizumab are beyond the ability of a resource poor setting. Therefore, the reduction or removal of risk factors, which we have just discussed, to

severe RSV disease might remain the most likely solution before a vaccine, which is affordable, safe and immunogenic becomes available. It would be difficult to mitigate or completely remove some of the risk factors e.g. having siblings in the household or time of birth relative to the RSV epidemic timing. However, parents can be sensitized about the benefits of exclusive breast feeding and reduction in the exposure to indoor smoke during the pre-natal period. This sensitization on a wide scale level e.g. country level may have significant benefits in reducing the number of not only RSV related severe disease but other respiratory infections.

2.5 State of RSV vaccine development

RSV is a high priority for vaccine development because of the high disease burden, community and hospitalized, that it causes in the very young children coupled with the fact that supportive treatment is the only alternative available to individuals who have developed severe disease. This supportive treatment remains very expensive in the developing world and would be far from affordable in many health care institutions. In the developed world, the cost associated with treatment of severe LRTI is significantly high [90, 145]. So if a vaccine becomes available, mathematical modelling has shown that vaccination of infants against RSV might be cost effective [132] and can potentially result in significant cost savings that can be directed to other projects.

One of the early RSV field vaccine trials was conducted in 1966 in California, USA [33]. The trial involved administering a formalin inactivated RSV vaccine to children aged between 4 months and 9 years. The investigators reported that in the subsequent RSV season when children were exposed to wild type virus there was

a significantly greater number of RSV infections admitted to hospital among the vaccine recipients compared to the non-vaccinated group in infants 4 months to 18 months. Clinically, the vaccinated group was hospitalized with more severe disease than the hospitalized non-vaccinated cases. However in older children, the response to natural infection seems unaltered by the vaccine. From the results, the vaccine proved to be non-protective. It appears that aside from failing to protect, the vaccine altered in some manner the host response to natural RSV infection. The exact mechanism under which the vaccine potentiated disease is not well known but it has been postulated that inadequate levels of serum neutralizing antibodies, lack of local immunity, immune complex deposition and excessive induction of a type 2 helper T-cell immune response [165] were responsible.

Different vaccine target populations have been proposed but the highest priority target population is children greater than 6 months old. These children are more likely to tolerate the vaccine with fewer adverse respiratory events since their immune system is more mature and they have lower levels of RSV specific maternal antibodies. Table 2.1 shows some of the other proposed target population groups for a vaccine population and the primary vaccine approaches. In the remaining part of this section, we will present a summary of the current status of some of the vaccines that are under development.

Table 2.1: Key target populations for an RSV vaccine. Table adopted from Anderson et al [5]

Target population	Key considerations	Primary vaccine approaches
0-6 months old infants	<p>Goal: prevent serious complications from infection</p> <p>Rationale: highest rate of hospitalizations</p> <p>Challenges: presence of maternal antibodies, immature immune system, susceptibility to RSV disease, history of the formalin inactivated RSV enhanced disease</p>	<p>1. Live attenuated RSV</p> <p>2. Live chimeric virus vectors</p> <p>3. Gene-based vectors 4. Potential for boosting sero-negative infants with sub-unit protein or particle-based vaccine after priming with live or gene-based vector vaccines</p>
Continued through the next page ...		

Target population	Key considerations	Primary vaccine approaches
6-24 months old children	<p>Goal: prevent serious complications from infection and reduce transmission to at-risk household contacts</p> <p>Rationale: \approx 50% of childhood hospitalizations occur after 6 months of age, maternal antibodies has waned, less susceptible to severe disease and more mature immune system than younger children, potential to decrease transmission to others</p> <p>Challenges: clinical endpoint may be more difficult to achieve than in neonates, history of formalin inactivated RSV enhanced disease</p>	<ol style="list-style-type: none"> 1. Gene based vectors 2. Live-attenuated RSV 3. Live chimeric virus vectors 4. Potential for immunizing RSV-seropositive children with subunit protein or particle-based vaccine or boosting sero-negative children after priming with live or gene-based vector vaccines

Continued through the next page ...

Target population	Key considerations	Primary vaccine approaches
Pregnant women or women of child bearing age	<p>Goal: increase passive antibody protection to foetus and prevent disease at most vulnerable age, block mother to infant transmission</p> <p>Rationale: high titre neutralizing antibody protects, can delay vaccination to older less vulnerable child with more mature immune system</p> <p>Challenges: having experienced multiple previous infections may limit response to vaccination, need for substantial increase in antibody levels to protect the infant, quantify the relationship between neutralizing antibody level and degree of protection</p>	<ol style="list-style-type: none"> 1. Subunit protein with standard adjuvants 2. Particle including virus-like particles (VLP) with standard adjuvants.

Continued through the next page ...

Target population	Key considerations	Primary vaccine approaches
Adults \geq 65 years	<p>Goal: protect from serious complications of infection</p> <p>Rationale: substantial RSV-associated disease in elderly population</p> <p>Challenges: having experienced multiple previous infections may limit response to vaccination, need to improve on protection provided by natural infection, difficult to diagnose and lack of clear indicators of the severity of RSV disease</p>	<ol style="list-style-type: none"> 1. Subunit protein with novel adjuvant 2. Particle include VLP with novel adjuvant 3. Gene-based vector with subunit protein or particle boost

rA2cp248/404/1030ΔSH is a cold-passaged, genetically engineered by reverse genetics live attenuated vaccine [114]. Its predecessor *cpts248/404* was administered in infants and caused significant nasal congestion that interfered with feeding and sleeping. Karron et al [114] have reported on a field study of *rA2cp248/404/1030ΔSH* and only 44% of infants who received two doses of the vaccine had detectable levels of antibody responses. The magnitude of vaccine virus shed was lower after the second dose than after the first indicating that an immune response capable of restricting viral replication had been induced. This surrogate for vaccine efficacy has further been supported by Wright et al [208]. Unlike the field vaccine of a live attenuated RSV vaccine in 1966 [33], enhanced disease was not observed when children and infants initially infected with vaccine virus were naturally infected with RSV [208, 114].

A phase 1/2a, randomized, double blind, placebo-controlled study to evaluate the safety, tolerability, immunogenicity of viral shedding of MEDI-559 (developed by MedImmune LLC and is a continuation of *rA2cp248/404/1030ΔSH*), a live attenuated intranasal vaccine against RSV in healthy RSV seronegative children 1- <24 months of age is currently underway [35]. The study comprises of two cohorts with the first cohort including RSV seronegative children between 5- <24 months while the second cohort is of infants 1- <3 months regardless of their baseline serostatus. The vaccine is given at a dosage of 0, 2 and 4 months. The primary outcome is the incidence of solicited symptoms after each dose through day 28 after each dose. The secondary outcome measure is the incidence and magnitude of MEDI-559 shedding at 7, 12 and 28 days after each dose. The study began in October 2008 and the expected date of completion was January 2012.

b/hPIV3/RSVF2 is a bovine PIV3 (parainfluenza virus) chimeric construct

that expresses the human PIV3 fusion (F), PIV3 hemagglutinin-neuraminidase and RSV-F proteins from the bovine PIV3 viral genome [194]. A phase 1/2a randomized, double-blind, placebo-controlled, dose-escalation clinical trial of the vaccine to evaluate the safety, tolerability immunogenicity and vaccine-like shedding against RSV in healthy 6-<24 months old and seronegative children and in 2 months old RSV immunity unscreened infants is currently ongoing. Doses of vaccine and placebo are given at 2, 4 and 6 months. The primary outcome measure is incidence of solicited symptoms from administration of study vaccine through 28 days following each dose while the secondary outcome is the incidence and magnitude of vaccine-like viral shedding at 7, 12 and 28 days after each dose [34]. However, in a double-blind placebo controlled trial involving 120 healthy adults aged between 18-40 years, the vaccine demonstrated that it was safe (produced no medically significant vaccine-related adverse events), did not boost RSV antibody titres from their baseline levels and was highly restricted in replication in seropositive adults.

Munoz et al [139] performed a randomized, double blind, placebo controlled study to determine the safety and immunogenicity of an RSV purified fusion protein (PFP-2) in women in the third trimester of pregnancy and their offspring. The vaccine was administered intramuscularly. The vaccine was safe, well tolerated and immunogenic in women in the third trimester of pregnancy. Seventy five percent of vaccine recipients had a response to PFP-2 by Western blot and 95% had a ≥ 4 fold rise in IgG ELISA Ab after immunization versus none in the placebo group. Transplacental transfer of anti-F IgG antibody was efficient and geometric mean concentrations of anti-F IgG antibody by ELISA were four fold higher in infants of vaccine recipients at birth, 2 and 6 months after delivery than in infants

of placebo ($p < 0.01$). Unfortunately, neutralizing response was generally quite low with $\approx 10\%$ of the vaccinated mothers showing a ≥ 4 fold rise in titre of neutralizing antibodies. Though potentially useful in adult populations, subunit vaccines such as PFP-2 may theoretically predispose RSV-naïve recipients for enhanced disease since they stimulate a bias toward a Th2 response [181]. Additionally, a subunit vaccine given to infants must be able to overcome any immunosuppressive effects of RSV specific maternally acquired antibodies. It is worth mentioning that a novel delivery method of RSV fusion protein vaccine in pregnant mothers is being developed by Novavax [64] using nanoparticles [105]. Nanoparticle delivery has the advantage of specifically binding to target cells and delivering high doses of the therapeutic contents. Target delivery is achieved by functionalizing the surface of the nanoparticles with proteins or small molecules that will bind to specific molecules on target cells while avoiding non-specific binding to other cells or tissues.

2.5.1 Challenges to RSV vaccine development

Young age at vaccination. The highest risk group for severe disease is in children less than six months of age [149, 150, 68, 82]. Thus an RSV vaccine would ideally be administered as soon as possible after birth. Healthy infants at this early stage of life have detectable levels of RSV specific maternal antibodies from their mothers [78, 152]. Therefore an RSV vaccine given in infancy would have to stimulate an acquired immune response in the presence of maternal antibodies. RSV vaccination will likely be delivered in the first six months of life at a time when other vaccines are given and therefore it will be of utmost important to

ensure that an RSV vaccine given at this age will not interfere with the safety and efficacy of routine childhood vaccines. In the event that the administration of an RSV vaccine through the standard immunizations schedule is not possible, then vaccination outside of the highest risk age group for severe disease would have to be investigated.

Possibility of severe disease. In the 1960's, a formalin inactivated RSV vaccine administered to infants and young children increased the risk of developing more severe disease following infection with the wild type virus [33] compared to children who did not receive the vaccine. This unfortunate clinical experience has led to heightened safety evaluation of candidate vaccines [114, 139]. Live attenuated vaccines are usually formulated to be given intranasally so as to elicit local immunity at the site of natural infection. Any live attenuated virus therefore would need to elicit immunity without causing inflammation of the respiratory tract.

In a recent review evaluating the global burden of RSV associated acute lower respiratory infections [145], it was estimated that 19.3-46.2 million new episodes of RSV associated ALRI occurred in children ≤ 5 years worldwide in 2005 with approximately 66,000-199,000 RSV related deaths with 99% of the deaths occurring in developing countries. So, the issue of developing a vaccine for the developing countries is key since they stand the highest chance of benefiting from such a vaccine. Clinical trials of candidate vaccines in these target populations are warranted but may be out of reach since they are expensive to carry out and time consuming. Hence progress of trials in low income countries requires the involvement of international organizations e.g. WHO (World Health Organization) or PATH (Program for Appropriate Technology in Health) to give their stamp of approval

and ensure that the vaccines are affordable in these populations.

To date, there remains no licenced vaccine for RSV and the need of preventing RSV associated illnesses in the population remains unmet. However a number of pragmatic recommendations have been proposed as the way forward by Anderson et al [5] as listed below:

- Develop good surveillance and disease burden data from the developing country setting to guide resource allocations decisions.
- Develop educational tools for patients, caregivers and much more importantly to government leaders in charge of making public health decisions to articulate the cost and potential benefits of an RSV vaccine.
- Develop RSV transmission mathematical models to illustrate how immunization may elicit both direct and indirect (vaccinating one target group may protect another population group) benefits.

In this project work, I have focused on the latter looking at the optimal age at which an RSV vaccine, should it become available, should be given. This involves the development of an age-structured mathematical model capturing the most important characteristics of RSV transmission, implementing vaccination in different age groups and then comparing the outcomes. However, the adoption of any single or combination of vaccination strategies would be subject to different factors including but not limited to: logistical, political goodwill and economic.

Chapter 3

Review of data on the epidemiology and transmission of RSV

3.1 Introduction

Since it was first isolated in 1957 [30], RSV has been identified as the most important viral cause lower respiratory tract infections in infants and children worldwide [145, 177]. Additionally, a number of studies have been carried out to enhance our understanding of the epidemiology of RSV. These studies include investigations of the role of maternal antibodies in protecting infants against disease and infection, the incidence of infection and disease in the community, and the duration of infection and immunity.

In this chapter, I will present some of the estimates of the epidemiological parameters important in describing the patterns of transmission of RSV in its host population. The reason for this review is two fold. Firstly, there have been recent advances in RSV vaccines. A recent trial of a recombinant live attenuated

vaccine demonstrated that the vaccine was well tolerated and immunogenic in the key target age group of 1-2 months [114]. For the evaluation of the vaccine efficacy, more vaccine trials are needed and there is an argument that they should be multi-country. There is therefore a requirement to establish the burden of the infection and disease in order to establish baseline denominator data for the selection of suitable sites that must meet the criteria of having this kind of detailed baseline data. Secondly, there has been an interest in describing the transmission of infectious diseases in communities using mathematical models and some have been used to evaluate the effectiveness of different control strategies. This has been true with RSV where a number of models have been published [201, 1, 205, 206]. The integrity of mathematical models mostly depend on the data that is available in order to parameterize them. This chapter will act as a reference point for which mathematical modellers working on RSV transmission models can get estimates of some of the most important parameters describing the transmission of RSV. In this chapter, I have presented a review of the available published data and I have pointed out gaps that may exist in the literature. A proper assessment requires a review of the available data on the duration of viral shedding, rate of infection, duration of immunity, duration of maternal antibody protection and the risk of development of disease and hospitalization given an infection.

3.2 Objectives

The objectives of the work presented in this chapter are to:

- estimate the epidemiological parameters relevant for the description of the transmission dynamics of RSV. The parameters considered are:
 1. duration of RSV specific maternal antibodies
 2. rate of infection
 3. rate of RSV disease
 4. duration of infection
 5. the duration of immunity
- determine the upper and the lower values for these parameters for inclusion in the sensitivity and uncertainty analysis in the work presented in Chapter 6.

3.3 Epidemiological parameters

3.3.1 Duration of RSV specific maternal antibodies

Children who are born to RSV seropositive mothers receive a similar level of RSV specific maternally derived antibodies. It is not clear what the presence of maternal antibodies (matAb) specific to RSV implies particularly when detected in serum, i.e. whether the protection is against disease [67, 118] or infection [154, 21]. Additionally, the age profile of the disease [28, 189, 148] would suggest that the duration of protection against infection is short, approximately between 1 and 3

months. The reason why the duration of protection is short is not well understood and it is plausible that it is due to inadequate matAb levels and/or matAb not being fully protective e.g. protective effect is strain specific [153].

A Brazilian study [37] investigating the seroprevalence of RSV antibodies in the community of a suburban population found out that the average duration of maternal antibodies was approximately 3.3 months (95% CI: 3-3.5). In this study, a serum sample was considered to be positive for maternal antibodies if it was above a cut-off threshold of approximately 1.7 log units. In a birth cohort study carried out in a rural Kenyan population in Kilifi [152], the average half life in days of RSV specific matAb for the seropositive population was 79 days (95% CI: 76-81) which is about 2.5 months, while the duration that an infant remained above the cut-off point for the detectable level of maternal antibody was 112 days (95% CI: 107-118) which is approximately 3.6 months. Another prospective study from Turkey monitored the concentration of maternal anti-RSV IgG antibodies in healthy newborns over the first six months of life [78]. Blood samples were taken at birth then at 1, 3 and 6 months of age. The mean antibody titre from birth to 1 month decreased by 38% and from 1 month to 3 months by 30%. At six months of age, four infants had positive RSV IgG which was interpreted as acquired infection. However, the use of anti-RSV IgG for the diagnosis of acute infection in young children may cause difficulties in interpretation since it is not possible to differentiate between maternal IgG and acquired IgG . When acquired infections were excluded from the analysis, the matAb positivity at 6 months of age was 2%. Considering the rapid drop in concentration and the high frequency of negativity (95%) at 3 months of age, it is proposed that these infants are susceptible to the infection before 3 months of age. Thus, it may be inferred that the mean duration

of protective RSV specific serum matAb is likely to be less than 3 months.

In another study, in the Netherlands [20], 45 children were enrolled at birth and followed for a period of six months in order to study the rate of decay of RSV specific matAb. The samples were assessed using the virus neutralization assay and competition ELISAs. At birth, neutralizing antibodies were present in the sera of all the 45 children. From the comparison of the virus neutralization serum antibody titres at birth with those found 3 months later, a mean half life of RSV specific matAb of 26 days was estimated. In another prospective study done in the UK [198], a group of newborn were followed for 12 months to measure the duration of protective RSV specific matAb. Initially, 100 infants and the mothers were enrolled but only 10 of these subjects were used in the study using radioimmunoprecipitation (RIFA) analysis. Blood samples were taken from each child at 3, 6 and 12 months after birth. The study reported a duration of maternal RSV specific antibodies to be 91.2 days. Table 3.1 gives a summary of the reported durations of decay of RSV specific maternal antibodies from these studies.

Table 3.1: Estimates of the duration of RSV specific maternal antibodies from published studies

Location	Sample design	Reason for sampling	No of subjects enrolled	Screening methods used	Age group	Mean duration	Refs
Kilifi, Kenya, 2002-2003	Infants recruited at birth. Homes were located in a health demographic survey area and within easy access	To monitor for infections and describe the age related serological changes	635 newborns	RSV specific IgG ELISA	0-6 months	112 days (3.6 months) (95% CI: 107-118)	[152]
Sao Paulo State, Brazil, 1990-1991	Random sampling of families from randomly selected administrative regions	Sera samples initially collected for the study of rubella prevalence	115 children	ELISA	0-6 months	3.3 months with range (3-3.5)	[37]
Continued through the next page ...							

Location	Sample design	Reason for sampling	No of subjects enrolled	Screening methods used	Age group	Mean duration	Refs
Bursa, Turkey, 2002	Random selection of pregnant women in a hospital setting. Children enrolled at birth then followed for 6 months	To determine the concentration of RSV specific maternal antibodies over the first 6 months of life	49 newborns	ELISA	0-6 months	3 months	[78]
The Netherlands, 1989-1991	Longitudinal study	Subjects initially participating in a Hepatitis B vaccination trial	45 children	Virus neutralization assay and competition ELISA	0-6 months	26 days ¹	[20]

Continued through the next page ...

Location	Sample design	Reason for sampling	No of subjects enrolled	Screening methods used	Age group	Mean duration	Refs
UK, 1983	Prospective study of a cohort of infants recruited at birth. At least one asthmatic member of the family was required	To determine the decay of RSV specific maternal antibody in infants	10 children	Radioimmuno precipitation assay, RIPA	0-1 year	91.2 days ¹	[198]

¹Estimated half life of RSV specific maternal antibodies

3.3.2 Rate of infection

The rate of infection is an important parameter estimating how the rate of spread of an infection changes with both age and time. It is formally defined as the frequency at which susceptible individuals contract the infection per unit time [10]. The rate of infection acting on a single susceptible in the population per unit time is referred to as the per capita rate of infection or the force of infection. Under the assumption of homogeneous mixing, the per capita rate of infection is an age-independent value. In our current RSV mathematical model developed in Chapter 6, we have relaxed the assumption of homogeneous mixing and therefore the force of infection is both a function of age and simulation time denoted by $\lambda(a, t)$ where a denotes the age class of the susceptible individuals and t is the simulation time in years. Additionally, the time dependent rate is seasonally forced using a cosinusoidal function in order to conform to the seasonality observed in peak RSV transmission. A number of community studies have estimated the rate of infection but most of these studies were hospital based with passive surveillance and hence they are likely to have underestimated the rate of infection in the community.

One study enrolled entire family members at the birth of a newborn and followed them for five years [68]. The members of the household were actively contacted weekly during the RSV disease season and nasal washes were obtained at the time of each visit. In this study, the rate of infection ranged from 82.6/100 cyo for children 13-24 months to 33.3/100 cyo for children in their fourth year of life. The results from this study are comparable with the results obtained from Henderson et al study [93] where they reported a crude estimate of the rate of infection as 53/100 cyo.

Hall et al [87] carried out another study in which they enrolled families on condition that each contained two or more children one of whom was less than a year of age. Enrolled families were visited every three to four days by a team of two nurses during the two months of the RSV season. Nose and throat specimens were taken of everyone in the household at the time of the visit and cultured for viral isolation. No serum samples were taken. All age groups had an appreciable attack rate as shown in Table 3.2 with a range of 16.8/100 in adults to 29.4 infections per 100 exposed individuals in infants. The crude attack rate recorded in this study may have been underestimated since some family members may have been infected before the initiation of the study. Weekly visits of the families were done after RSV was confirmed to be in the population after a child was hospitalized with RSV. Additionally, infected individuals may have had negative culture through technical difficulty in handling the the relatively labile virus.

In another study by Henderson et al [93], study subjects attended a child development center evaluating the effects of educational intervention on the psychosocial and cognitive development of normal children. A daily observation of their respiratory health was made. Upper respiratory tract cultures for virus were taken every two weeks irrespective of symptoms and at the onset of each illness. Fifty percent of the enrolled children (39 out of 78 children) were followed through a minimum of 4 exposures with an exposure defined as an yearly seasonal outbreak. As shown in Table 3.2, the rate of infection during the first exposure (assuming that the first exposure corresponds to primary infection) was 98/100. The rate of infection during the second exposure was 75/100 and that for the third exposure was 65/100. From this study, the highest rate was recorded in children in their first years of life with a reduction in that rate in the second and third years of life.

Given that the study was carried out in a day-care setting, it represents a special epidemiological place where the conditions are best suited to the spread of a respiratory infection requiring close contact for transmission. Although it is not clear how these circumstances differ from those found in a family with young children, it is expected that the conditions of exposure in a day-care center approach the maximum encountered by most children. The reported results are higher when compared to those obtained from the family study by Hall et al [87].

Nokes et al carried out a study in a rural population in coastal Kenya where a birth cohort study was recruited and followed for one year [150]. During the study, weekly visits were undertaken and a nasal washing collected when clinical symptoms of acute respiratory infection (ARI) were observed. Passive surveillance was done via parental referral to Kilifi District Hospital. The incidence rate (IR) of RSV infection was reported to be about 42.8 cases per 100 cyo. The low value of the reported figure relative to estimates from other cohort studies [93, 68] may be due to differences in study design, population settings and methods of determination of RSV infection. For example, in the study by Glezen et al [68], weekly samples were collected irrespective of symptoms and infections were confirmed through virus isolation or a four fold increase or greater in serum neutralizing antibodies while in this study, nasal samples were obtained only when subjects presented themselves with clinical symptoms of ARI. Both of these reasons would have resulted in increased sensitivity of case detection.

Borrero et al [19] also carried out a study in Cali Colombia. A cohort of 340 children were followed from birth for a period of 17 months. The children were chosen from families of low socioeconomic strata. Active weekly visits were made to the homes of the participants and the children with signs of ARI were referred to

a health center from where the following samples were taken; nasal aspirates, nasal swabs, blood and urine samples. The crude rate of infection was reported as 19.97 cases per 100 cyo. The age specific rates of infection are shown in Table 3.2. The reported estimates are close to other studies reported previously. However there are some reasons to believe that the incidence reported may be underestimated. The number of episodes counted depends in part on the number of weeks a child was in the study and whether a child visited the clinic when ill. It would be difficult to know what proportion of children with ARI who were not referred by the home visitor to the clinic although the authors suggest that it was very small. It is also reported that 10% to 15% of all those referred to the health center did not present themselves. Whether they went to another clinic or not can not be fully ascertained. Therefore the true rate of infection would be higher than reported here. RSV was also reported to be the most common cause of ARI with the rate of infection slightly higher for children less than 12 months of age.

Weber et al [203] used both hospital surveillance and community-based study to obtain information about the spread of RSV in the compounds in which infected children lived. Upon diagnosis of a case of RSV in the hospital, the household of this so called index case was visited as soon as possible, and a regime of twice weekly follow-up for six weeks initiated. During this follow-up, children under five years of age that were examined and found to exhibit signs of acute respiratory infection had a nasal aspirate collected. Table 3.2 shows the reported rates of infection where the lowest rate was recorded in children between 2-3 years (16/100 cyo) while the highest was in children between 0-1 and 4-5 years (33/100 cyo). The reported rate of infection may be under estimated because compounds of households of the hospitalized index child, usually the infant, were visited only

after the hospitalization i.e. infection in the household may have preceded disease in the index child.

Table 3.2 shows a summary of the reported estimates of the rate of infection from the studies described. Where applicable, the rate of infection is categorized by the age class of the participant and the history of previous infections.

Table 3.2: Estimates of the rate of infection from published studies

Location	Sample design	Case ascertainment	Age class	N ¹	Rate/100 person years at risk			Refs
					Crude	First infection	Re-infection	
Houston, USA	Entire family enrolled at the birth of a newborn	Active weekly home visits during RSV season. Nasal washes obtained irrespective of symptoms and virus cultured. Blood collected to detect infection by seroconversion or 4-fold rise in titre using neutralization assay.	0-12 m 13-24 m 25-36 m 37-48 m 49-60 m Total	92	68.8 82.6 46.2 33.3 50.0 62.9	68 97.1 100 - - 74.4	- 75.9 45.3 33.3 50.0 53.0	[68]
Continued through the next page ...								

Location	Sample design	Case ascertainment	Age class	N ¹	Rate/100 person years at risk			Refs
					Crude	First infection	Re-infection	
Rochester, USA	Entire family enrolled	Active visits during RSV epidemic every 3 to 4 days over the two month study. Nose and throat specimens were obtained irrespective of symptoms for virus isolation.	< 1yr	34	29.4	-	45.4	[87]
			1-<2yrs	7	28.6	-	0.0	
			2-<5yrs	34	26.4	-	16.6	
			5-<17yrs	48	18.7	-	21.0	
			17-45	55	16.8	-	33.3	
		Total	178	21.9	-	27.0		
Cali, Colombia	Birth cohort of 340 children. Women enrolled when attending pre-natal clinic	Active weekly visits were made at the home of the participants. Sample taken for children with signs of ARI, nasal aspirate, nasal swab, blood and urine sample	0-5 m	340	19.14	-	-	[19]
			6-11 m		25.32	-	-	
			12-17 m		12.48	-	-	
			Total		19.97	-	-	

Continued through the next page ...

Location	Sample design	Case ascertainment	Age class	N ¹	Rate/100 person years at risk			Refs
					Crude	First infection	Re-infection	
Chapel Hill, USA	Children were attending a development center assessing the effects of educational intervention on psychosocial and cognitive development	Active daily observations. Upper respiratory cultures were taken every 2 weeks and at the onset of symptoms. 50% of the children were followed over a minimum of 4 years	0-1yr	61	98.4	-	-	[93]
			1-2yrs	47	74.5	-	-	
			2-3yrs	26	65.4	-	-	

¹N represents the sample size

3.3.3 Rate of RSV associated LRTI

All of the studies reported in section 3.3.2 report on the incidence of RSV infection. However, in this section we will report on studies that have reported the rate of severe RSV disease in order to establish the burden of disease caused by RSV by which comparison can be made with the burden of other infectious diseases competing for health resources. This is potentially beneficial in offering policy makers a platform from which they can best make health decisions when faced with different conflicting health needs.

In a study carried out in Brazil by Suttmoller et al [193], active surveillance was done among children in two low income populations in Rio de Janeiro. There was also inpatient and outpatient data collected over the study period of 3 years from January 1987 to December 1989. Weekly home visits were made by the health care workers and nasal aspirates were taken from children with symptoms of lower respiratory infection (LRI) . The rate of RSV disease per 100 children at risk varied in the three groups i.e. community, inpatient and outpatient, with the inpatient recording the highest at 40 and the lowest rate was reported in the community at 22. The age-specific risk of disease is as reported in Table 3.3. During the 3 study periods, RSV was observed to be seasonal.

Karron et al [113] carried out a study in the northern part of the United States in Alaskan native children less than three years old and enrolled them into the study when admitted with ARI at the YK delta regional hospital. Each subject had a nasal aspirate taken for virus isolation and antigen studies. To calculate the incidence of severe disease, suspected nosocomial infections were excluded from the analysis. Forty six percent of all the ARI hospitalizations and 31% of all

hospitalizations were due to RSV. Nineteen percent of all children admitted with RSV were readmitted. Given that surveillance in this study was passive and that they relied on children who presented themselves to the hospital for admission, it may under estimate the actual community risk of disease.

Robertson et al [171] reported on study in four developing countries namely; Mozambique, Nigeria, Indonesia and South Africa in order to determine the age-specific incidence of disease and seasonality of RSV associated respiratory infection in children less than 5 years. For the active community surveillance sites, i.e. Nigeria and Indonesia, there were weekly visits to households with children aged less than 5 years and samples were collected from children with LRI. In South Africa and Mozambique, there was passive surveillance through a health care provider and children who presented to the hospital with symptoms of LRI were enrolled for the study. The incidence of disease from the different locations is as recorded in Table 3.3. The recorded risks for the four different areas were different with an order of magnitude of 10 between the lowest and the highest values. A factor that may have led to these differences may include the sheltering of infants for the first 40 days after birth in Nigeria and Indonesia. In the South African study site, rates of health care utilization may have been high but there was a documented drop in clinic attendance during the study due to a shortage of drugs.

Another study was carried out in a rural coastal population in Kenya to report on the incidence of RSV infection and disease [150]. Infants were recruited at birth or within two weeks of birth and active weekly household surveillance was done during the epidemic period and monthly otherwise. During the visits, a nasal washing was collected if there were signs of acute respiratory illness. Blood samples were taken at birth and at an interval of approximately 3 months until the study

completion. From this study, the crude incidence of disease after RSV infection was reported as 37/100 child years at risk with the highest rate being reported in children aged 3-5 months (36/100) and the lowest risk reported in the children aged 6-8 months (33/100). A more recent re-analysis of these data has been done taking account of denominator cases due to serologically determined infection [155]. Following the re-analysis, the crude incidence of disease following RSV infection has been reported as 27/100 with the highest risk reported in children 0-2 months (44/100) and the lowest in children aged 18-23 months (18.6/100). See Table 3.3

Berman et al [15] enrolled children less than 15 years to a two year ambulatory study in Cali Colombia. A passive surveillance system was set up in 5 health centers in Cali and children were enrolled in the study if signs of acute lower respiratory tract infections (LRTI) were observed. A blood sample and nasal swab were taken for culture and a throat sample was taken if there was a clinical diagnosis of pharyngitis. The crude incidence of LRTI was reported as 0.6/100 children at risk.

Djelantik et al [43] have reported the incidence of RSV associated disease in a community in Lombok, Indonesia taking part in a Haemophilus influenzae type b vaccine. As part of the study, RSV testing of children less than 2 years hospitalized with severe LRTI was performed. Nasopharyngeal wash samples were taken and tested for RSV positivity using a rapid enzyme immunoassay. For children less than 1 year of age, the incidence of disease was reported as 0.8/100 child years at risk. However, when accounting for untested cases (assuming that tested and untested cases had the same proportion of RSV positivity) the corrected estimate is 2.5/100. One of the limitation of the study is the inability to obtain specimens of all the children, including most of those who died or were discharged without

a sample being taken. For this reason, the reported estimates will likely be an underestimate of the true risk of severe disease with RSV.

Chan et al [29] reported on a retrospective study conducted at a teaching hospital in Hong Kong of children <5 years hospitalized with severe RSV disease. Patients with laboratory confirmed RSV infection during the 5 years study period (Jan 1993- Dec 1997) were identified through records of the virus laboratory. RSV was detected from the patients' nasopharyngeal aspirates, bronchoalveolar and endotracheal aspirates by direct immunofluorescent staining and in parallel by virus isolation. The risk of RSV hospitalization with severe disease was recorded as 0.25/100 in children < 5 years. Table 3.3 shows a summary of the incidence of severe disease.

Table 3.3: Estimates of the rate of RSV associated severe disease from published studies

59

Location	Sample design	Case ascertainment	Age class	N ³	Rate of severe RSV disease/100 cy			Refs
					Crude	First in- fection	Re- infection	
Rio de Janeiro, Brazil	All children < 5yrs enrolled following baseline census of two low income communities	Active weekly home visits. Nasal aspirated from children with LRI were collected.	0-5 m	262 ⁵	18.1	-	-	[193]
			6-11 m		9.1	-	-	
			12-23 m		5.2	-	-	
			24-35 m		2.4	-	-	
			36-59m		1.5	-	-	
Alaska, USA	Hospital based surveillance of children less than 3 years	Passive surveillance of Alaskan native children between October 1993 to September 1996 . Nasal aspirates were obtained for viral isolation and antigen studies	<1 yr	32 ⁴	93/94 ¹ -5.3	-	-	[113]
				152 ⁴	94/95 ¹ -24.9	-	-	
				95 ⁴	95/96 ¹ -16.4	-	-	

Continued through the next page ...

Location	Sample design	Case ascertainment	Age class	N ³	Rate of severe RSV disease/100 cy			Refs	
					Crude	First infection	Re-infection		
09	Bandung, Indonesia	Community surveillance of children less than 5 years	Weekly household visits. Children with signs of LRI were escorted to a clinic where a nasopharyngeal specimen as collected by a physician	<1 yr < 5 yrs	1420 ⁵ 4.1 3.4	- -	- -	[171]	
	Manhica, Mozambique	Hospital based surveillance of children less than 5 years	Passive surveillance of outpatients < 1 yr old. Nasal specimen were taken for RSV antigen test using ELISA.	<1 year	-	3.0	-	-	[171]
	Ibadan, Nigeria	Community surveillance of children less than 5 years	Weekly household visits by research nurses. Nasal specimen collected from children with signs of LRI and were tested for RSV antigen using ELISA	<1 yr <5 yrs	- 11.6 9.4	- -	- -	[171]	

Continued through the next page ...

Location	Sample design	Case ascertainment	Age class	N ³	Rate of severe RSV disease/100 cy			Refs
					Crude	First infection	Re-infection	
Agincourt, South Africa ²	Hospital based surveillance at the clinics of children less than 5 years	Passive surveillance at 6 primary healthcare clinics of patients who presented with severe LRI. Nasal specimen was taken and RSV detected using ELISA.	<1 yr <5 yrs	-	1.5 0.9	- -	- -	[171]
Houston, USA	Entire family enrolled at the birth of a newborn	Active weekly home visits during RSV season. Nasal washes obtained irrespective of symptoms and virus cultured. Blood collected to detect infection by seroconversion or 4-fold rise in titre using neutralization assay.	0-12 m 13-24 m 25-36 m 37-48 m 49-60 m Total	125 ⁵	22.4 13.0 10.8 7.7 0 14.5	21.6 5.9 0 - - 18.1	- 19.0 10.9 7.7 0 11.4	[68]

Continued through the next page . . .

Location	Sample design	Case ascertainment	Age class	N ³	Rate of severe RSV disease/100 cy			Refs
					Crude	First infection	Re-infection	
Kilifi, Kenya	Birth cohort. Children enrolled at birth or before 2 weeks after birth	Active weekly visits during RSV seasons and monthly otherwise. Nasal washings were taken when signs of ARI were observed. Passive surveillance was through the Kilifi District Hospital (KDH). Blood samples were taken at birth and approximately 3-month intervals	0-2m	635 ⁵		44.2	18.5	[155]
			3-5m			42.9	18.5	
			6-8m			23.3	18.5	
			9-11m			22.0	18.5	
			12-17			23.9	14.7	
			18-23			18.6	14.7	
			≥24m			18.8	19.6	
			Total			27.0	16.6	

¹The risk of severe disease is given by the epidemic year i.e. 93/94, 94/95 and 95/96.

²This study only reports the risk of severe LRI. This may explain why it records the lowest estimates compared to the rest of the studies which record the risk of LRI.

³N represents the sample size

⁴Sample size shown grouped by age or year of participation

⁵Sample size not age/group specific

3.3.4 Duration of infectiousness

Three durations are important to distinguish in the study of the epidemiology of RSV, namely, the duration between infection and infectiousness (latent period), the duration between infection and the onset of clinical symptoms (incubation period) and the period between the beginning of infectiousness and the cessation of infectiousness [156] often referred to as the infectious period. All the three periods will vary between individuals and in particular, there will be a distributed period between infection with virus and the beginning of infectiousness for which it is difficult to have information about in an observational field study. In our investigations of transmission dynamics of RSV, we are interested in the latent and infectious periods. A recent systematic review has reported that the median incubation period for RSV is 4.4 days (95% CI 3.9-4.9) [119, 112]. However, what is observed is usually the start and cessation of symptoms from which to infer estimates of interest. However, the use of clinical symptoms to guide estimates of latent and infectious periods can be significantly biased if there is either shedding before or after symptoms start. There is therefore a need to carry out studies that collect samples irrespective of clinical symptoms. However, in estimating the shedding duration, a number of studies have accounted for the importance of the incubation period since the study design is such that there was nasal sample collection irrespective of the clinical symptoms. For studies that have reported the duration of shedding based on hospitalized individuals, it is likely that they underestimate the duration.

In an investigation of the transmission dynamics you are interested in the latent and infectious period. However, what is observed is usually the start of symptoms

(and cessation) from which to infer estimates of the parameters of interest. However, use of clinical symptoms to guide estimates of latent and infectious period can be significantly biased if there is significant shedding before (and after) symptoms start (stop). Therefore it is important 9a) to collect samples irrespective of symptoms- as done in Patrick's study.

In this section however, we will concern ourselves with reviewing the infectious period of RSV which is defined as the period between the beginning of viral shedding, into the nasopharynx, to the cessation of viral shedding. An individual is considered to be shedding RSV if the virus or virus specific components can be detected in nasal secretions through diagnostic techniques ranging from culture, Immunofluorescence Antibody Test (IFAT) to Polymerase Chain Reaction (PCR). In order to estimate the duration of infectiousness, we have assumed that the beginning of viral shedding corresponds to the start of infectiousness and the cessation of viral shedding corresponds to loss of infectiousness and therefore we are estimating the duration of viral shedding as a correlate for the infectious period distribution (see the discussion).

In Turku, Finland, [199] a study was undertaken to investigate the shedding of infectious virus during acute infection with RSV. The study population consisted of children hospitalized at Turku University Central Hospital during an RSV season from March 1989 to February 1990. During the period of hospitalization, nasal aspirates were taken for a study on the comparison of methods for RSV diagnosis. One or more follow up specimens were collected from 41 randomly selected patients with proven RSV infection for the shedding duration study. Followed up children had a mean age of approximately 8.6 months. From this study, 40-60% of the patients ceased to shed the virus 8-10 days after admission to hospital.

In a study done in Rochester, USA, [89] adults were challenged intranasally with a safety tested pool of RSV at increasing intervals over a 26 months period. Challenges were administered six times at 2, 4, 8, 14, 20 and 26 months. Nasal washes were obtained before each challenge and daily for the next two weeks. A serum sample was also taken at the time of natural infection. From the study, fifteen adults were identified as having natural RSV infection with a mean duration of viral shedding at 4.7 days with a range of 1-8 days. Of the total infections resulting from the challenge, the average duration of viral shedding was 3.4 days with a range of 1 to 7 days. After the first reinfection, the average duration of shedding was 4.6 days compared to an average of 1.7 days for the subsequent reinfections. Given that this study was done in adults who would most likely have had experienced their primary infection, then the results presented in this study would possibly be an under estimate of the duration of shedding in primary infecteds. Furthermore, the amount of virus inoculated in the participants during the experimental study may differ from what one would get during a natural infection and perhaps this would influence the shedding duration.

In another study conducted in Rochester, USA during the period from 1975 to 1995, a total of 2960 healthy adults (18 to 60 years) were prospectively evaluated for respiratory virus infections [88]. Surveillance was done during the 5 to 6 months when RSV was actively identified to be present in the community. Evaluation of all subjects was conducted 2 to 3 times per week and a nasal wash sample was taken alongside a physical examination to determine if the illness was symptomatic. Acute RSV infection was identified in 7% of these adults. The average duration of viral shedding among 118 infected individuals was 3.9 days with a range of 1 to 17 days. Shedding was detected for less than 7 days in all but 8 of these individuals.

In a different study, infants hospitalized with respiratory syncytial virus infection were studied to describe the quantitative shedding pattern and the duration of shedding [85, 80]. Nasal washes for viral culture were obtained from hospitalized subjects with acute respiratory tract disease during a two month period when RSV was epidemic in the community. The samples were collected as soon as possible after admission and thereafter every one to three days during the period of the infants hospitalization. Some infants who appeared to be shedding the virus at the time of discharge were followed at home where nurses obtained nasal wash specimens. For the patients who were followed up for the entire RSV shedding period, the mean duration of shedding was reported to be 6.7 days with a range of 1 to 21 days. However, infants with lower respiratory tract disease shed for a significantly longer period, 8.4 days, compared to those with upper respiratory tract infection, mean 1.4 days.

In another study, in Houston, USA, [59] children and adults were studied within a family study. Families were enrolled for prospective study of viral respiratory infections with the birth of a new infant. Nurses took viral specimens (nasal washes or throat swabs) from all children during the home visits scheduled weekly or biweekly depending on the season. This study reported over 70% of cultures positive up to 8 days post illness onset. Majority of the individuals with primary infections and reinfections shed the virus for approximately 8 days.

In a study done in a coastal population in Kenya [158], 635 infants were recruited at birth and intensively monitored for acute respiratory infections. A sub sample of 70 households were enrolled into a family study. All family members were actively monitored for ARI through weekly household visits during epidemic periods and monthly otherwise. Nasal washings were collected from infants and

children younger than 15 years experiencing episodes of acute respiratory illness. The overall duration of recovery irrespective of infection history, age and severity of illness was 4.5 days (95% CI: 4.0 to 5.3 days). The community study done in Kilifi additionally suggests that the duration of shedding between children who had never been infected and those with prior history of infection differed by approximately 40% i.e. 4.9 days (CI:4.1-5.8) for primary infection compared to 4.1 days (CI:3.3-5.1) for secondary infections. For children presenting to the research clinic and whose illness history was obtained, the duration of shedding among children with no previous history of infection was 8.2 days (95% CI: 6.5-10.3) compared to 7.0 days (95% CI: 5.5-8.8) of those with prior history of infection.

A human experimental infection model has been done that included 35 healthy adult (18-45 years) volunteers [41]. Subjects were admitted to a quarantine unit for 13 days and were observed for at least 1 day before RSV inoculation, which occurred the second day after admission. Nasal washes were obtained on the day of admission and twice daily on day 1 up to day 12. Pulmonary tests were also performed daily on all volunteers. The mean duration of viral shedding was 7.4 days, ± 2.5 days, as assayed by qPCR. Duration of shedding was lower as assessed by quantitative culture (3.6 ± 1.1 days) and spin-enhanced culture (5.3 ± 1.4 days).

Hall et al [87] reported on a family study designed to examine intrafamily spread of RSV infections and their associated illnesses. They enrolled families on the basis of including two or more children. Families were visited every three to four days by a team of two nurses during the RSV epidemic season. Nose and throat specimens were taken of everyone in the household at the time of visit. The duration of documented shedding was between 1 to 36 days with a mean of 3.4 days. The mean duration of shedding for children less than 16 years of age was 3.9

days while for those over 16 years the mean duration was 1.6 days. Children less than two years of age had shedding for significantly longer periods with a mean of 9 days.

Munywoki et al have recently carried out a household study within a rural coastal population in Kenya to determine the duration of viral shedding [142]. A household based prospective cohort study was set up with a recruitment target of 50 RSV naive infants and their household members. Trained field assistants made household visits every 3 to 4 days to collect nasal samples irrespective of symptoms from all the household members. Samples were taken using nasopharyngeal flocked swabs and follow up was done during the 2009-2010 RSV season. An individual RSV episode was defined as the period within which an individual provides specimen that were PCR positive for the same infecting group not more than 14 days separating any two positive samples. The estimated mean duration of viral shedding based on the midpoint estimate was 11.2 days (95% CI: 10.1-12.3). The duration of shedding differed between those who had symptomatic infections and those with no symptoms: 13.5 vs 7.8 days. The age-specific durations are shown in Table 3.4 which also gives a summary of the previously discussed studies.

Table 3.4: Estimates of the duration of shedding

Location	Sampling design	Number enrolled	Screening methods used	Severity of disease	Age groups	Prior history of infection	Duration of shedding			Ref
							Mean	Mode	Median	
Turku, Finland	Hospitalized children with ARD. ¹ Two or more specimens were taken during hospitalization.	41 children	NPA ² collected for viral culture and detection of RSV antigens using TR-FIA ³	Children were hospitalized with ARD	Age range: 2 wks - 3.75 yrs			9days Range: 8-10 days		[199]
Continued through the next page ...										

Location	Sampling design	Number enrolled	Screening methods used	Severity of disease	Age groups	Prior history of infection	Duration of shedding			Ref
							Mean	Mode	Median	
Rochester, NY	Adults who had acquired natural RSV. Nasal samples were collected daily for 2 weeks after challenge.	15 adults	Nasal specimens collected and virus isolation was done using immunofluorescent testing	6: moderately ill, nasal congestion 5: moderately severe with fever 4: mild URTI ⁴	Adults		4.7 days Range: 1-8			[89]

Continued through the next page . . .

Continued through the next page . . .

Location	Sampling design	Number enrolled	Screening methods used	Severity of disease	Age groups	Prior history of infection	Duration of shedding			Ref
							Mean	Mode	Median	
Kilifi, Kenya	Entire family enrolled	50 households. 179 individuals with RSV were included in this analysis	Active household visits every 3-4 days. Screening irrespective of symptoms. Nasal specimens were taken	Ranged from no symptoms to symptomatic	<1yr 1-4yrs 5-14yrs 15-39yrs ≥40yrs		18.0d 11.8d 9.1d 8.4d 11.2d	-	-	[142]

¹Acute Respiratory Disease

²Nasopharyngeal Aspirate

³Time-resolved Fluoroimmunoassay

⁴Upper respiratory tract infection

⁵Lower respiratory infection

⁶Immunofluorescent antibody test

⁷Lower respiratory tract infection

3.3.5 Duration of immunity

The components of immunity to RSV and its durability are not well understood. Children have been shown to develop antibody to both F (Fusion) and G (attachment) proteins but the role of these antibodies in protecting against infection or disease in humans is not well defined [198]. Additionally, although young children are known to be infected repeatedly, even during successive annual epidemics [93, 3], the duration of immunity and influencing factors remains relatively unknown. In this section, I have presented a review of the studies that have sought to estimate the duration of protective immunity against RSV infection.

In a study carried out in Kilifi, Kenya, enrolled children in the birth cohort were monitored through active household visits, weekly during the epidemic periods and otherwise every 4 weeks and passively through referral to a research outpatient clinic at the Kilifi District hospital [150, 184]. Infections were considered as separate episodes if a positive result was determined ≥ 14 days after a previous positive. Molecular relatedness of RSV causing primary and repeat infections, by phylogenetic analysis, in 12 infants from the birth cohort was used to provide insight into the duration of RSV immunity. The total observation period spanned 16 months and 2 epidemics. Results from this study have indicated that re-infections with RSV occurred not only during the first year of life but also during the same epidemic and that re-infections can arise with the same variant within 7-9 months and in the same group but with different variants within 2-4 months. The average duration of immunity irrespective of the re-infecting variant, whether homologous or heterologous to the primary infection variant, is approximately 6 months. This is clearly influenced by the seasonal nature of the virus

occurrence [3]. A more recent re-analysis of the data from this cohort evaluating the genetic relatedness of infecting and reinfecting RSV strains demonstrated that the mean interval between infection and re-infection was 365 days (range: 21-699 days) [3]. Of the 55 reinfections, approximately 13% occurred within the same epidemic period with a mean interval of 30 days (range: 21-56 days) while 80% occurred in either the same or consecutive epidemic as the first infection while the remaining 20% occurred in non-consecutive epidemics. Using data derived from this cohort [155], it has been demonstrated that, after infection, immunity to reinfection appears to last up to 6 months and is 60%-70% partially effective.

Mufson et al [137] did a study to investigate how often second infections are with a virus of the alternate subgroup or alternatively with the same subgroup. This study included 13 children with acute upper or lower respiratory tract infection for whom RSV was isolated from pharyngeal swab specimen on two occasions at least 9 months apart. These children were part of an RSV surveillance carried out in Huntington in the United States of America (USA). Overall, second infections with the homologous subgroup were detected as often as second infections with the alternate subgroup and all the re-infections were at least 9 month later (at most 26 months later) with a mean immunity period of approximately 17 months.

To understand the duration of immunity against respiratory syncytial virus, Hall et al [89] experimentally infected 15 healthy individuals who had acquired natural RSV infection. The subjects were intranasally challenged with RSV at increasing intervals of time at 2, 4, 8, 14, 20 and 26 months after natural infection with a similar strain group (A). After each challenge, the subjects were evaluated daily for 2 weeks by physical examination and nasal washes were also obtained. Thirty three percent of the individuals were infected after the first challenge, 2

months after natural infection, while subsequent challenge produced infection in 25% - 30% of the subjects. 73% of the subjects were re-infected one or more times during the 26 months after natural infection and about 50% had three or more infections during the same period. Although this work does not report a mean duration of immunity, it seems that protection is far from solid and is potentially of short duration given that there were some participants who got infected 2 months after the natural infection.

3.4 Summary and discussion

In this chapter, I have presented a summary of the various studies reporting on some of the characteristics of the natural history of RSV infection that are relevant for its epidemiological study. One of the most important risk factors associated with the risk of hospitalization early in life i.e. mostly within the first 2 months of life, is the level of maternal antibody. The importance of maternal antibodies to RSV in infants has previously been a subject of much discussion [65, 154, 50, 51] and it has been shown that children with elevated levels of maternal antibody to RSV are protected from infection longer compared to children with low levels of antibodies [154, 67]. One of the risk factors associated with decreased placental transfer is the level of maternal antibody concentration [49] in cord blood. There exists seasonal variation in the RSV specific matAb titre. Antibody titres are highest following an RSV season [179]. Higher titres have also reported from mothers who had other children at home, perhaps due to the frequent contacts with the children who might be infected [87, 88]. However, the duration of the protective level of maternal antibodies to RSV remains poorly understood. As shown in Table 3.1, different studies have reported conflicting estimates. Most of the reported estimates have been about the duration of detectable levels of RSV specific matAb rather than the duration of the protective effect. Furthermore it has not been established whether the presence of maternal antibodies in sera is an indication of whether an individual is protected from infection or not. One of the studies suggests a linear relationship between the level of maternal antibody and the age at first infection [67]. Additionally, the authors suggest that for maternal antibody in blood to offer protection against infection, the anti-RSV antibodies need to diffuse

through to the lining of the respiratory tract at the alveolar and bronchiolar level to provide protection against serious lower respiratory tract disease. It is not known what levels of maternal antibody in blood is required for such a diffusion to occur hence offering protection. Due to the lack of a credible duration of protection of maternal antibodies and a huge variation in the reported estimates, we estimated the duration of RSV specific matAb by fitting the mathematical model describing the transmission dynamics of RSV to hospitalization data as explained in Chapter 6. Additionally, we included the parameter estimating the duration of protective level of maternal antibody in the sensitivity and uncertainty analysis in order to evaluate what effect a variation in the point estimate has on the number of children hospitalized with RSV and how that influences the outcome of vaccination against RSV.

The rate of infection on the other hand is an important parameter determining how quick the virus is able to spread. There are huge variations in the estimates of the rate of infection from the different studies reported here. It is difficult to tell what level of variation is due to methodological differences or due to intrinsic variation in the study populations. Factors that affect the estimation, accuracy and the interpretation of disease and infection incidence data are discussed. The experimental design is an important factor. Cross sectional studies require recruiting newborn and upper censoring at the upper age limit while a cohort is vulnerable to bias in exposure risk. In order to offset bias arising from temporal variation, the recruitment should be done throughout the year. Another important factor is the method with which cases are determined. Active community case ascertainment may encourage individuals to participate while passive case ascertainment is influenced by care seeking behaviour. It has previous been reported

in the Gambia [203] a decrease in incidence of disease with increase in cost of travelling to the hospital and this has also been reported by Nokes et al [148]. The method of sample collection also seem to differ across different studies with known differential sensitivity. For example, in a recent study investigating the sensitivity and specificity of real time multiplexed PCR (M-PCR) and immunofluorescence in the detection of respiratory viruses, the authors found out that nasopharyngeal flock swabbing was superior to nasal wash for the detection of viruses by M-PCR (sensitivity, 89.6% versus 79.2%; $p=0.0043$) although inferiority of the nasal wash was not observed when immunofluorescence was used to detect the presence of the virus [141]. Hence differences in laboratory diagnosis methods will also influence the outcome. Given that the published estimates of the rate of infection give widely varying estimates of the force of infection which are confounded by some of the factors discussed above, we will estimate the force of infection in the model using the relatively new but well established *social contact hypothesis* [196].

Studies that investigate the duration of shedding are quite few and especially those cases that are not hospitalized. Even fewer of these studies are reported from developing countries. Out of the eight studies reported here, 5 were hospital based (2 for children and 3 for adults) and only 4 community studies. Community studies where sampling of participants is done irrespective of the symptoms provide the most accurate estimates of the duration of shedding. From the studies recorded in Table 3.4, it seems that estimates of the duration of infection recorded from hospital admissions are higher than for community studies. This may be as a result of the fact that most hospital admissions are severe and they also represent the narrow age group of children less than two years of age. It has been shown that greater severity of RSV infection within in-patients results in increased duration

of shedding [85]. Data from the family studies by Munywoki et al [140] and Hall et al [87] provide the best estimates given the design of the study. Participants were sampled every 2 days during the RSV epidemic season and irrespective of their infection status. This design approach reduces the bias that is brought about by both right and left censoring of data. We have therefore used 9 days as the mean duration of shedding for primary infecteds in the mathematical model developed in Chapter 6. For the re-infections, Okiro et al [158] suggests that children with a history of RSV infection have a 40% increased rate of recovery from infection i.e. shorter duration of viral shedding and therefore duration of shedding given a history of infection was taken to be approximately 4 days [158]. This estimate is similar to that recorded by Hall et al [88] during yearly community surveillance of 2,960 adults between 1975 and 1995. The rate of recovery from infection in the model is then assumed to be the reciprocal of the estimated duration of shedding. here are a number of factors that influence the comparability of studies reported in section 3.3.4. Sensitivity of the detection of viral shedding (a measure of infectiousness) is likely to be determined by the amount of virus shed, which is expected to increase and then decrease as the infection progresses [41, 82]. Hall et al [82] have shown that as the days post infection increase, subsequent specimens tend to have lower viral load. From this observation, we can conclude that sensitivity of detection will be variable within the infectious period of an individual and participants sampled during early or late stages of infection are less likely to be classified as shedders.

There are however a number of factors that influence the reliability and comparability of the studies reporting on the duration of infection. Sensitivity of the detection of viral shedding (a measure of infectiousness) is likely to be determined

by the amount of virus shed, which is expected to increase and then decrease as the infection progresses [41, 82]. Hall et al [82] have shown that as the days post infection increase, subsequent specimens tend to have lower viral load. From this observation, we can conclude that sensitivity of detection will be variable within the infectious period of an individual and participants sampled during early or late stages of infection are less likely to be classified as shedders.

Another likely bias is the identification of failures i.e. recovered individuals who have stopped viral shedding. Different studies have used different criteria to try and distinguish between prolonged shedding and unique shedding periods. For example, Okiro et al [156, 158] defined a failure as a single negative sample preceded by a positive sample. So, given the nature of error in biological measurement, and especially at low level shedding, it is possible that children testing negative on a single day might subsequently turn positive on the next day. To minimize this bias in sampling, one would define a failure as two or three subsequent negatives.

Another factor influencing the reliability of the estimates is left and right censoring both of which will tend to under estimate the shedding duration. For most studies, it is difficult to determine the start time (left censoring) of shedding. This bias is more enhanced in hospital based studies where sampling will begin only after hospitalization or presentation to the hospital [199, 85] by which most individuals will have already started shedding. Studies that take samples irrespective of symptoms provide more accurate information although that also depends on the interval of sampling. Daily sampling would be desirable but due to the invasive nature of the sampling procedure, this would be difficult to implement. On the other hand, right censoring (difficulty in determining the time of viral shedding cessation) is more common in hospital based studies if participants are discharged

while still shedding. Some studies have tried to minimize this bias by following participants at home after discharge [85, 80].

Other factors that have been shown to influence the duration of shedding is previous exposure to the virus [158] where it has been reported that the duration of shedding is reduced by up to 40% in the secondary infecteds compared to primary infections. However studies have indicated that there is no correlation between age and the duration of shedding [158, 156, 85]. Infants with lower respiratory infection have been reported to shed the virus for a significantly longer period (mean 8.4 days) than those with upper respiratory tract infection (mean 1.4 days) [85]. Given this observation, it is therefore possible that estimates of the duration of viral shedding reported from children with lower respiratory tract infections are an overestimate of the mean viral shedding of the general population.

The extent to which the studies can be compared is also dependent on the nasal sampling method and the laboratory diagnosis method. Nasopharyngeal aspirates (NPA) and nasal washes (NW) have been the gold standard for the diagnosis of respiratory viruses [81] but in a recent study in children with mild respiratory illness, it has been demonstrated that nasopharyngeal flocked swabbing (NFS) is superior to NW collection for the detection of viruses by real-time multiplexed PCR (M-PCR) (sensitivity, 89.6% versus 79.2%; $p=0.0043$) [141]. However, NFS collection was non-inferior to NW collection in the detection of RSV by immunofluorescence antibody test (IFAT)

The duration of acquired immunity to infection is expected to have a significant contribution to RSV epidemiology. The duration of immunity determines how frequent individuals become susceptible to re-infection and this may have a potential bearing on the seasonality of the infection and hence disease. There have not been

many studies documenting the duration of immunity to RSV. The data collected from the Kilifi birth cohort [184, 3] presents the best estimate for use in modelling studies. We have therefore taken the duration of immunity after infection to be approximately 6 months. The rate of loss of immunity is calculated as the reciprocal of this duration. Due to the variation in the estimates, the parameter estimates reported in this chapter and which are used in the mathematical model developed in chapter 6 will be part of the sensitivity and uncertainty analysis to evaluate what likely effect a change in their value will have on selected model outputs e.g. on the number of predicted hospitalization before and after vaccination.

Chapter 4

Social contact patterns relevant for the spread of RSV

4.1 Introduction

The spread of respiratory infectious diseases depend on the social mixing patterns that bring individuals into contact sufficient to facilitate transmission. Since mixing is intimately related to age, the age structure of contacts i.e. within and between age classes is of key importance in determining the pattern of spread [48, 136]. Age specific contact rates have therefore been useful in modelling the transmission of respiratory infectious diseases and potential impact of vaccine administration on transmission pattern [164, 110]. In the case of Respiratory Syncytial Virus (RSV), a major respiratory pathogen, transmission is via fomites and large respiratory droplets [86, 83], which for effective transmission, require close contacts.

Estimation of transmission parameters has previously been done using case no-

tifications or age cross-sectional serological data [10, 54] and more recently through socio-demographic data [60, 40] and self reporting diaries [136, 196, 48]. Such data have been of considerable use in the past but have significant limitations: i) estimation of the next generation matrix (NGM) from age related force of infection has an identifiability problem, ii) steady states of the endemic disease equilibrium are required and iii) the method is best suited for to infections for which age-prevalence continues to rise over a wide age range, otherwise inference is considerably reduces. This method has worked well for childhood infections such as measles, mumps, rubella and chicken pox [180, 9, 10]. Contact data on the other hand yields estimates of who mixes with whom but the disease specific probability of transmission following a contact has to be indirectly estimated.

A number of diary contact studies have been done and their investigations can be broadly categorized as those that constitute either short conversation without physical touch or a conversation involving physical touch [136]. It has been shown that the pattern of contacts within and between age groups identified from data on conversations is qualitatively matched by that of physical contacts, [136, 110, 40], the latter however being less frequent and more likely to result in complete records. An important finding in diary based studies is that mixing is assortative with age [136, 48, 110] i.e. people of similar age-group tend to mix more often than those in different age-groups. Observations between age-groups contacts more likely represent the interaction between individuals of parental age and young children and referred to as cross-generational mixing.

Such data on social contact mixing is largely lacking within the developing country setting. In fact, only one such study has been done in Africa by Johstone et al [110] and it was recently published in 2011. In general, the areas in which these

data were collected were either non-representative of our setting [136, 100, 95] or there were design issues, such as small convenient sample sets [16, 48] and therefore it is not clear to what extent the result can be generalized to our geographical setting. To address this lack of empirical data from our setting, we undertook a population based, prospective diary survey of epidemiologically relevant social self-reported contacts. A random sample of individuals from the registers of the KHDSS were selected by age class and asked to keep a diary (record of physical contacts occurring with a day). The resulting data was used to extract daily contact rates by which to define the age-specific mixing pattern.

Given the scarcity of contact data in different populations, the costly nature and the difficulty in carrying out a population wide contact study, we also developed a computation approach to derive mixing patterns from routinely collected socio-demographic data. In particular, we focus on constructing a synthetic contact matrix from available demographic data on households. This kind of synthetic contact matrix generation has recently been independently developed for a number of European countries with a notable agreement between the synthetic mixing matrices and the contact diary data generated by the POLYMOD study [60]. The proposed method of generating the synthetic matrix is general and can be easily adopted in other regions where household demographic data is readily available.

The two contact matrices developed in this chapter i.e. diary mixing and synthetic mixing data have been used in the mathematical model developed in Chapter 6 to model the transmission dynamics of RSV and to explore the benefit of different vaccination strategies.

4.2 Objectives

The objectives of the work developed in this chapter are to estimate the age-specific contact rates in the KHDSS from:

- conventional prospective self-reported contacts from diary data
- a synthetic mixing matrix generated from demographic household data

4.3 Methods

4.3.1 Prospective diary survey

4.3.1.1 Study design

The study is a cross-sectional survey with randomized sampling by age-group. Six age-groups were identified: infants (< 1 year old), pre-school (1-5 years old), primary (6-14 years old), secondary (15-20 years old), adults (20-49 years old) and elderly (> 50 years). A contact is defined as an interaction between two individuals involving some form of close physical contact e.g. handshaking, kissing or embracing.

4.3.1.2 Study site

The study was conducted in the northern part of the KHDSS of Kilifi District which is a coastal District in Kenya, with the Kilifi District Hospital as the reference point. Participants were drawn from 5 locations that are traversed by the main Kilifi-Malindi highway i.e. Kilifi Township, Tezo, Ngerenya, Roka and Matsangoni as shown in Figure 4.1.

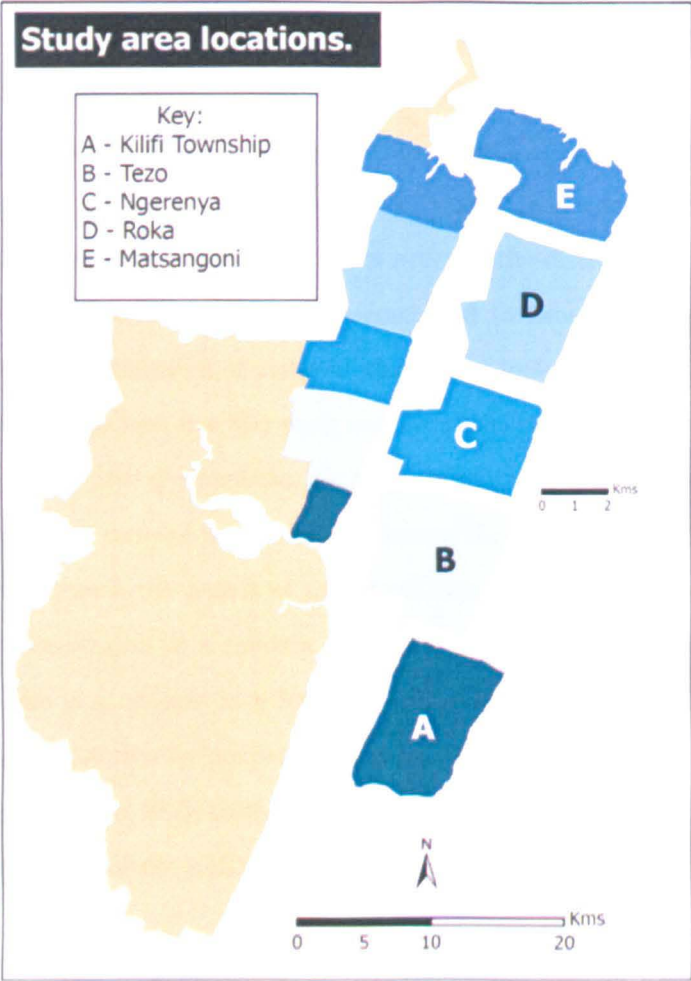


Figure 4.1: KHDSS map with the regions labelled A, B, C, D and E showing the study locations within the KHDSS

The KHDSS is located on the Indian Ocean coast of Kenya and was established in 2000 as a record of births, pregnancies, migration events and deaths and is maintained by 4-monthly household visits. The study area was selected to capture the population from which hospitalizations are observed at the Kilifi District Hospital. It has a total population of approximately 262,000 as at March 2011 (See Figure 4.2 for the population structure) living within an area of approximately 900km² with 49% of the population aged less than 15 years. It is worth highlighting at this point that there exists clear differences between the distribution of male and females above 20 years i.e. fewer males compared to females. This is possibly due to the high outmigration of males at this age to look for work outside of the KHDSS. Most of the men live and work outside of the KHDSS and more particularly in the nearby city of Mombasa. As they maintain their family in rural Kilifi within the KHDSS, some of them consider themselves resident too. Some mothers as well move between the homes of their husbands and they therefore don't fall within the classification of a resident. These two issues highlight the challenge of defining who is a resident in a frequently migrating population. The average crude out-migration rate for the period between January 2006 and December 2010 has been estimated at 88.5/1000 pyo (persons year of observation). For a more detailed description of the KHDSS, please see the work by Scott et al [182].

4.3.1.3 Study population

The individuals who met the following conditions were included in the study:

- residents of the KHDSS
- participants able to give a written informed consent or their parents/guardians

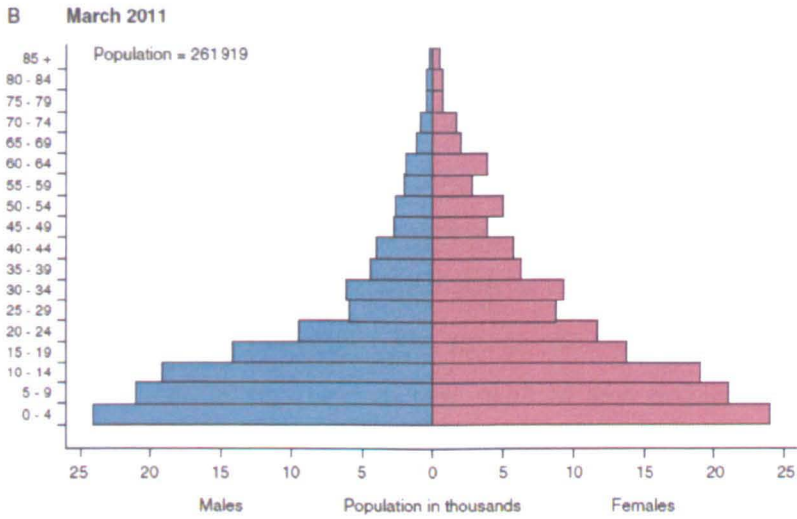


Figure 4.2: Population pyramid of the KHDSS in 2011. Adapted from Scott JAG et al [182]

if below the age of 18 years

Individuals were excluded from the study if:

- they refused to give informed consent.
- they were planning to move out of the KHDSS within three months from the study start date.

4.3.1.4 Sample size calculation

An age structured mathematical model of RSV transmission was used to conduct a sensitivity analysis of the influence of variation in the age-specific contact rates on the estimated age-specific force of infection (a study outcome influential to the predicted impact of intervention). In the absence of suitable data for our proposed

study population, we adopted the age-specific rates of physical contacts defined in the POLYMOD study of UK residents [136]. The daily contact rates shown in Table B.1 in Appendix B was scaled by a factor q in order to estimate the per capita rate of infection per age group according to the social contact hypothesis [196]. The contact rates for the six age-groups are scaled by a multiplicative factor of $\pm 20\%$ (range 0.8-1.2) and the resulting matrix is then fitted to RSV inpatient data from the Kilifi District Hospital. The age specific force of infection is then plotted at equilibrium i.e. at the final time point as seen in Figure 4.3. The variability observed in the per-capita rate of infection gives an indication of the relative importance of the contacts between the age groups involved. For example, Figure 4.3 first panel shows the force of infection for each of the multiplicative factors as shown in the legend, when the contact rate between infants and infants is increased/reduced by $\pm 20\%$. Considerable variability is observed which is an indication that the contacts between infants and other infants is important in determining the pattern of spread of the infection in the population of interest. In the final panel, i.e. the one on the 6x6 location, little variability is observed when contacts between adults older than 50 years are changed within the same limits. From Figure 4.3 we can see that the variation in contact rates within and between three groups, namely, infants, preschoolers and adults have more influence on the force of infection estimates than those of the other groups. For example, a 20% change in the contact rate of infants resulted in 0.69% change in the force of infection in infants and 9.13% in adults, whereas a 20% change for secondary school children resulted in 0.033% change in the force of infection for this group and 0.0018% for younger age groups. This implies that minimum sample size estimates should be based on achieving adequate precision of contact rate estimates in these

three age groups.

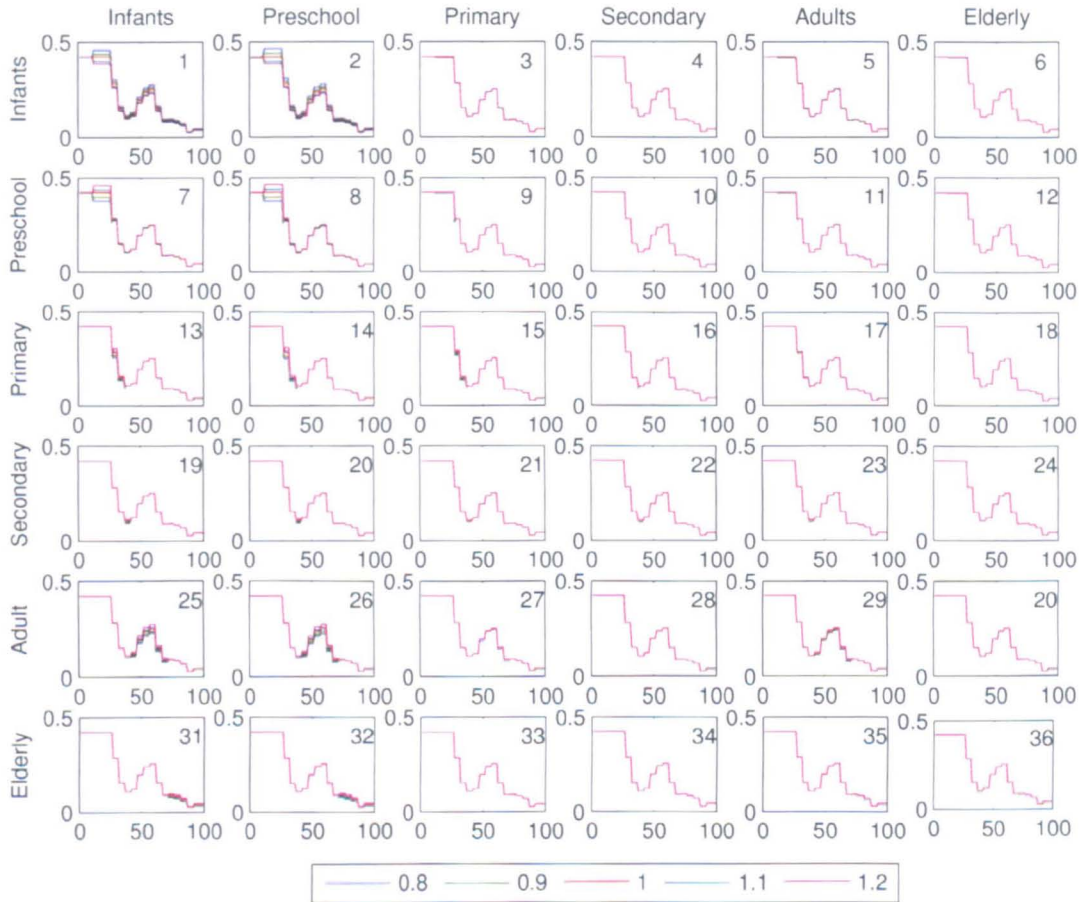


Figure 4.3: Sensitivity measured as the variability observed in the force of infection when the daily contact rate between two classes e.g. infants and preschoolers, is varied between $\pm 20\%$. The lines in each panel show the age-specific force of infection as predicted by the model output at equilibrium.

Assuming a desired precision (i.e. 95% confidence limit) of $\pm 20\%$ for an estimated contact rate, and using an estimate of the contact rate variation (standard deviation ≈ 13) based on a school study diary recently undertaken in the KHDSS

but not yet published (SSC#1716, n=177), we estimate using standard methods a required sample size of 150 per age group or a total of 900 for the six age classes. To account for possible non-response and errors in completion of the diary, this number was increased by 20% to give a final sample size of 1080 individuals.

4.3.1.5 Sampling procedure

The study participants were drawn, in equal number, from each of the 5 locations shown in Figure 4.1 using records available in the KHDSS stratified into the six age classes. Recruitment was also staggered over a six month period spanning a period of social mobility related to farming practice, with equal sampling effort by month. A random sample of 180 individuals was therefore selected from each age group.

The selected participants were followed up for consenting after explaining the requirements of the study. During consenting, participants were informed that they will be keeping the diary on one random day of the week. Each participant was requested to randomly pick one card out of seven, with each card labelled with a different day of the week. The participants were visited one day prior to the assigned day by a suitably trained field worker who explained how the diary was to be filled-in and also collected basic demographic data about the participant. The diary was collected the day following the recording day allocated to the participant. The field worker then conducted an exit interview to assess the participant's views on the diary and also asked a set of structured questions that will be used to determine the commonness of contacts with each individual and assist in validating the data collected in the diary.

Random sampling of the entire study area was performed on a monthly basis

for five months. This translates to 180 diaries per month with each of the 5 field workers coordinating the fieldwork expected to consent and supervise the filling of 9 diaries per week. This monthly staggering allowed us to account for heterogeneities in age-specific contact rates that occur before, during and after the normal planting season and between rural and town locations.

4.3.1.6 Tool specification

Two types of diaries we proposed. Text diaries and pictorial diaries. As is common in social science studies, a pilot study was conducted among 50 participants selected at random from the study area. The purpose was to assess the suitability of the diaries and make improvements, identify a suitable method of reminding participants to complete the diaries before rolling out the main study and also to formulate a validation method using a standard exit questionnaire. From the Focus Group Discussions (FGDs), we adopted a diary with both text and pictorial descriptions of the age classes for all study participants, see the diary sample in Appendix A.1.

The diary design is such that the participants will only record the information necessary for use in the age-structured RSV transmission model developed in Chapter 6. To minimize the effort required to fill out the data in the diary and to maintain the correctness of the data entered, we used a simple diary where only the age class of the person contacted and the frequency of the contact was entered. The diary was designed as a booklet with instructions on the first page on how to fill it in and the remaining five pages were reserved for recording contacts. In total, the diary had a capacity to record contacts with 75 different individuals. However, there was no limit to the number of physical contacts recorded with the

same individual in the final column of Appendix A.1. The duration or intensity and location of contact were not included in the diary design. The exit interview ascertained the frequency of contacts for each individual recorded in the diary to try and ensure the correctness of the data. The diaries were also translated to and from Swahili and Giriama (local dialect) from English. A participant was then required to select a diary in one of the languages provided.

Codes representing household members were written on the diary by the participant with the help of the fieldworker prior to handing over the diary to the participant. A household was defined as people who eat from the same kitchen.

4.3.1.7 Community engagement

We set up a number of focus group discussions (FGDs) during piloting to identify, for example, how to best shadow those unable to self-complete the diary and how to best fill in the diary including prompting reminders through wrist watches or mobile phones. The FGDs were convened in the community with the help of the Community Advice for Specific Study Teams (CAST) at the KEMRI-Wellcome Trust Research Programme, Kilifi. The CAST team is responsible for advising and assisting study teams on the best approaches and practices when conducting community based studies.

4.3.1.8 Ethical considerations

The study underwent local internal review by the KEMRI/Wellcome Trust Programme Scientific Coordinating Committee prior to submission to KEMRI Scientific Steering Committee. Subsequent ethical approval for the study was sought and granted by the KEMRI/National Ethical Review Committee (KEMRI/RES/7/3/1)

as well as the Biomedical and Social Ethics Review Committee of the University of Warwick (134-07-2011).

4.3.1.9 Diary allocation and keeping procedures

Participants were expected to keep the diary for a day, and this was generally expected to begin when the participants woke up and end when they went to bed [48]. A field worker visited the participant a day prior to the randomized day of keeping the diary and explained the procedure for making diary entries. An at risk event which was to be recorded was a direct physical contact. A direct physical contact was defined as having occurred when two or more people touch one another e.g. handshake, sleeping together, embracing and kissing. Sharing of objects and talking without touching was not to be recorded as a contact. Repeated contact with the same individual was to be recorded in the same row as a tally in the last column of the diary. See Appendix A. Participants were requested to record information about the age class of each contact person. The diaries had six age classes which were chosen to so as to represent the schooling pattern in the population up to the secondary school and then adults and elderly in the final two age classes. The six age classes are infants (< 1 year), preschool (1-5 years), primary school (6-15 years), secondary school (16-19 years), adults (20-49 years) and elderly (>50 years). These age classes were represented in the diary as both text and pictorial for ease of identification by the participants as shown in Figure A.1. During the first visit, the fieldworker also administered a questionnaire recording the sociodemographic information about the participant and that of the family members. These data included the age of the participant, the occupation, the gender, level of completed education, the administrative location and the com-

position of the family. Participants were requested to record their contacts as they happened. In instances where this was not possible, they were asked to fill the diary at the next convenient time and were supplied with a reminder table, a separate piece of paper to keep a record of these contacts before transferring them to the diary. Participants were provided with alarm watches that were programmed to go off at hourly intervals so as to remind them when to fill the diaries.

Individuals who could neither read nor write were assisted in filing their diaries by shadows. The shadowing process was designed to be as covert as possible so that participants did not modify their normal behaviour. Participant and shadow were expected to meet once hourly with alarm serving as a reminder and the participant was supposed to tell the shadow the people with whom they have had a physical contact in the previous one hour i.e. the age class and the number of times contacted. Shadowing of preschoolers and infants was done by the caregiver at home. Consented adults who can neither read nor write selected an appropriate person to shadow them. School going children < 9 years old were shadowed by two people, one in the household and the other one while at school. At the household, the caregiver shadowed and while at school, the class teacher was asked to assist the child to record their contacts. Both the caregiver and the class teacher gave informed consent or assent to allow participation in the study.

4.3.1.10 Data capture, management and analysis

Data capture and storage Each field worker was appropriately trained to explain the requirements of filling in the diary to the participants with an emphasis on ensuring the participant or the shadow understood how the diary was to be filled. Participants were also encouraged to fill in the diary once the wrist watch prompt

went off or whenever possible so as to mitigate recall bias. Data from the diaries was entered into a customized database in FileMaker® on the KEMRI-Wellcome Trust Research Programme servers in Kilifi.

Data analysis The average number of contacts per person was computed from the tally of the total contacts i.e. all contacts occurring within and outside of the household. In addition, the mean number of contacts stratified by age-group, smoothed and corrected for reciprocity were calculated as explained below.

Generation of the contact rate matrix Let Y_{ij} be the number of unique individuals (respondents) in age class j that a participants in age class i contacts. Then Y_{ij} has the observed values $y_{ij,t}$ and $t = 1, 2, \dots, T_i$ where T_i is the number of participants in age class i . Therefore, Y_{ij} can be expresses as the total number of contacts as:

$$Y_{ij} = \sum_{t=1}^{T_i} y_{ij,t} \quad (4.1)$$

To calculate the mean number of the unique individuals contacted per day per person, we divide each row by the number of participants in each age class denoted by T_i . Let us denote the mean number of unique individuals contacted per day by $\mu_{ij} = E(Y_{ij})$ such that:

$$\mu_{ij} = \frac{1}{T_i} Y_{ij,t} \quad (4.2)$$

where μ_{ij} denotes the mean number of unique individuals contacted per day a participant in age group i has with individuals in age group j .

Contact surface smoothing Contact surface smoothing was performed using a smoothing spline. The smoothing was performed in order to explore the non-

linear structure of the data and to find the relationship between age and the unique number of contacts. The idea behind the smoothing spline is to combine a measure of the smoothness of a function and how well it fits the data. In this report, we have used a Bayesian formulation of the smoothing spline as reported by Rue et al [176]. A smoothing spline formulation involving a Bayesian model with a Gaussian prior is given as:

$$f(x) = \theta_1 + \theta_2 x + b^{-\frac{1}{2}} F(x) \quad (4.3)$$

$$y_i = f(x) + \epsilon \quad (4.4)$$

where $x \in [0, 1]$, $i = 1, \dots, n$ and $\epsilon \sim \mathcal{N}(0, \sigma^2)$. The coefficients θ_1 and θ_2 can be fixed and unknown or random, $b > 0$ is a precision parameter and

$$F(x) = \int_0^1 (x-t)_+ dW(t) \quad (4.5)$$

where $(\cdot)_+ = \max(0, \cdot)$. $F(x)$ is a one-fold integrated Weiner process and the solution of the stochastic differential equation (SDE) shown in Eqn (4.6). The smoothing spline can therefore be found by solving a SDE and a Bayesian model.

$$\frac{d^2 f(x)}{dx^2} = \frac{dW(x)}{dx} \quad (4.6)$$

The Bayesian hierarchical model introduces a latent variable in addition to the observations and the hyperparameters such that

$$\text{Observations: } y|f \sim \pi(y|f) \quad (4.7)$$

$$\text{Latent variable: } f|\theta \sim \pi(f|\theta) \quad (4.8)$$

$$\text{Parameters: } \theta \sim \pi(\theta) \quad (4.9)$$

The distribution of latent variable and parameter $\pi(f|\theta)$ and $\pi(\theta)$ represent $f(x)$ and θ before the observations are done and are referred to as the priors. Rue et al [176] have developed a computer package in R, for approximate Bayesian inference using integrated nested laplace approximation (INLA). The package is designed to handle Gaussian models such as the Bayesian formulation of smoothing splines and computes the 95% confidence interval from the posterior densities without using MCMC methods.

Correction for reciprocity The major assumption behind this correction is that at the population level, symmetry of total contacts must hold. That is, the total number of contacts that individuals in age group i make with respondents in age group j must be equal to the total number of contacts that individuals in age group j make with participants in age group i . We will compute the total number of contacts, at the population level, between age group i and j and denote it by T_{ij} such that:

$$T_{ij} = \mu_{ij}P_i \quad (4.10)$$

where P_i is the number of people in the population in age class i . If we assume reciprocity of total contacts holds at the population level, then

$$T_{ij} = T_{ji} = \mu_{ji} P_j \quad (4.11)$$

However, $T_{ij} \neq T_{ji}$ and some of the reasons why the discrepancy may exist is because of participants failing to record all their contacts, recording more contacts than they actually had, participants contacting people outside of the population and bias in sampling. To correct the lack of reciprocity, we calculate the mean number of contacts between age group i and j from both T_{ij} and T_{ji} as shown in Eqn (4.12).

$$M_{ij} = \frac{T_{ij} + T_{ji}}{2}. \quad (4.12)$$

The adjusted mean number of contacts per day per person after correction for reciprocity is given by

$$\mu_{ij}^R = \frac{M_{ij}}{P_i} \quad (4.13)$$

and we refer to this matrix as the diary contact matrix.

4.3.2 Synthetic mixing matrix

The diary contact matrix in the previous section requires one to conduct a survey in order to estimate the mean number of contacts per person per day. Such surveys are not available for all populations because such collection of empirical contact data on large scale is both difficult and expensive. In this section, we will present an alternative approach by constructing a synthetic contact matrix. The main advantage with this method is that it is general and can be easily used for

regions without contact surveys data but with the necessary socio-demographic data available.

We developed the synthetic matrix with the starting assumption that the WAIFW matrix can be thought to be composed of three sets of contacts, namely: household contacts, school contacts and general/other contacts. In order to construct the overall contact matrix, the three setting specific matrices were built and then linearly combined. In generating the synthetic matrix, a contact is assumed to have occurred if two or more people share a physical environment.

Household contacts These are contacts that happen within the household setting. A household in our setting is defined as an establishment where people share and eat from the same kitchen. The contacts that occur within this setting can be thought to represent a stable and easily quantifiable compartment of population mixing [136]. They are also characterized by some common features of intense and regular interactions. To determine the entries of this matrix, we obtained household data from the KHDSS that was collected between September 2010 and January 2011. These data include the location i.e. latitude and longitude, of each household, the number of people and the age of each individual in the household. Using this set of data, we generate a household mixing matrix assuming that individuals within each household mix homogeneously but by age. We started by putting people in yearly age classes up to 76 years of age and people older than 76 years were put into the last age class. Whenever we encountered two people from the same household, we added a counter to the corresponding intersection of their age classes. Since contacts are reciprocal in nature, the corresponding intersection of their ages was updated as well. All self-contacts were disregarded.

We repeated the process for all the individuals in the population in each and every household. This matrix is represented as H_{ij} and represents the total number of contacts in the population occurring in the households. Since H_{ij} accounts for the total number of contacts accounting for reciprocity, then it relates to the mean number of contacts C_H as shown in Eqn (4.14)

$$C_{H_{ij}} = \frac{H_{ij}}{w_j w_i} \quad (4.14)$$

where w_i, w_j are the population sizes in age classes i and j .

School contacts In general, contacts within a school setting have been reported to occur within individuals of predominantly the same age group [133]. We have assumed that children will be in school between the age of 5 years and 20 years. To populate this school mixing matrix, we equated the elements of the main diagonal from age 5 to 20 years with the maximum value from the household contact matrix. Additionally, we equated the two parallel diagonals on either side of the main diagonal to half of the value of the main diagonal of the household matrix. This matrix is represented as C_S .

General/other contacts A homogeneous contact matrix represents all other contacts in the population outside of the household and the school. The homogeneous mixing matrix assumes that the mixing rate between age groups is independent of the number of individuals within the age groups. This matrix is represented by C_{HS} .

To relate the three mixing matrices to the WAIFW matrix, we linearly combine

them as shown in Eqn (4.15)

$$WAIFW = q_H C_H + q_S C_S + q_{HS} C_{HS} \quad (4.15)$$

where q_H , q_S and q_{HS} are disease specific infectivity parameters and are estimated by fitting the the age structured RSV model developed in Chapter 6 to RSV hospitalization data. However, in this chapter we will present the results only of the household component of the synthetic matrix since the other two component matrices are based on assumptions rather than generated from data. Additionally, the WAIFW matrix generated from the linear combination shown in Eqn (4.15) is presented in Chapter 6.

4.4 Results

4.4.1 Diary survey results

A total of 1138 individuals were randomized for participation in the study. Fifty five percent (623 individuals) consented to participation while 515 refused to participate. Table 4.1 shows the reasons for non-consent in the study and by the study locations as shown in Figure 4.1. Of the 623 diaries given, 606 diaries were collected and only 571 were completed and were used in the analysis presented. The mean age of the participants is 23 years (SD 22).

Table 4.1: Reasons for non-consent

Reason for non-consent	A	B	C	D	E	Total
Hostile	0	0	0	0	1	1
Silent	1	1	0	1	1	4
Not interested	28	23	6	9	6	72
Absent	14	8	2	6	19	49
Migrated	30	5	4	4	14	57
Temporarily away	27	11	10	17	19	84
Parent failed to give consent	21	3	7	13	2	46
Untraceable	23	3	0	2	1	29
Withdrawal	6	7	0	2	2	17
Dropped from study	9	0	0	0	2	11
Died	1	1	1	1	1	5
Unknown	49	17	21	36	17	140
Total	209	79	51	91	85	515

A total of 27,395 physical contacts were recorded with 11,498 unique contact persons. Figure 4.4 shows the distribution of total contacts which are skewed to

the right with majority of the participants recording contacts of frequency 1. The baseline characteristics of the 571 study participants who completed the diaries are shown in Table 4.2. Of the 27,395 total physical contacts recorded, over 70% of these contacts were reported during weekdays. The mean number of people contacted per day per participant was associated with the participant’s age class with the lowest mean number of contacts reported among infants (15.16) and the highest among the secondary (22.8) followed by the primary (21.9) school ages classes. Figure 4.5 shows the distribution of the contacts that are with unique

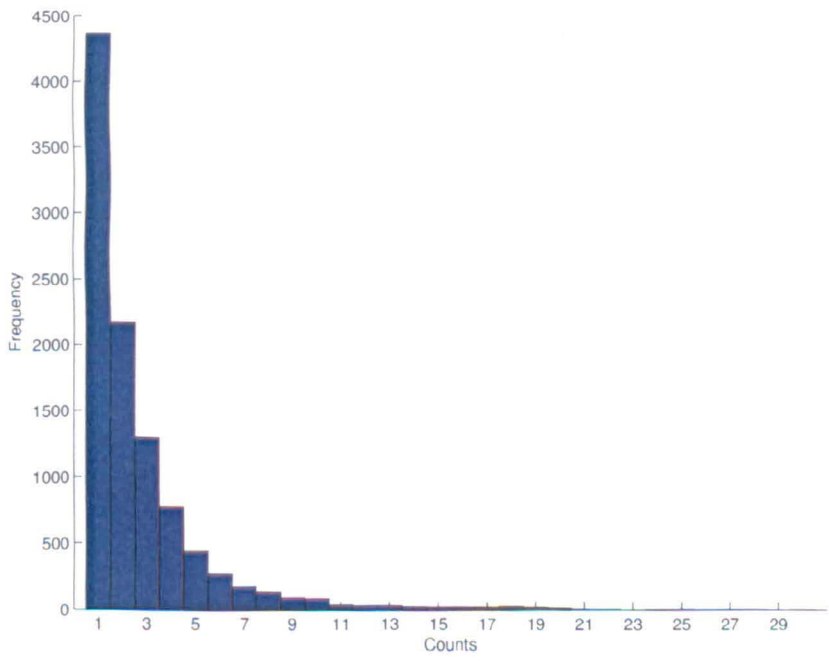


Figure 4.4: Distribution of contacts. Note that the contacts are skewed to the right with majority of the participants recording contacts of frequency 1.

individuals. From the figure, we can observe that the median number of individuals contacted is 17 per person per day. The distribution is skewed to the right with 3

participants recording the highest number of unique people met in a day at 68.

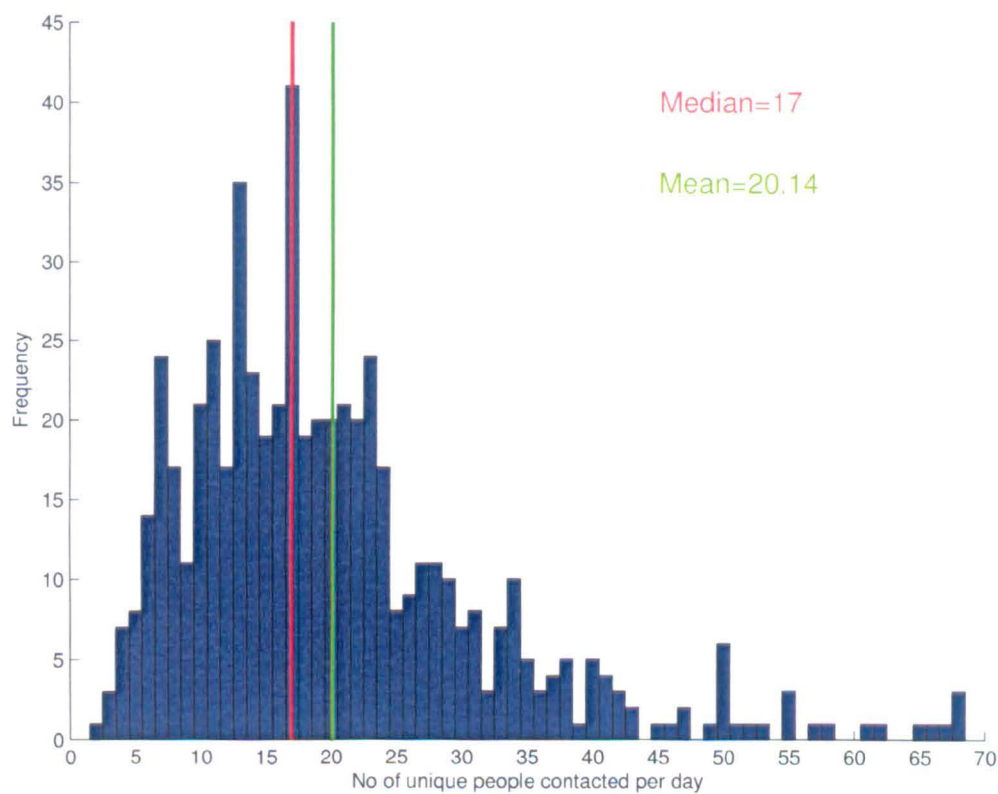


Figure 4.5: Shows the distribution of contacts that are with unique individuals with a median (red) of 17 and a mean (green) of 20.14 per person per day

4.4.1.1 Age related mixing pattern

Figure 4.6 shows the mean number of people contacted (solid grey lines) with the smoothing splines (black solid lines) and the 95% CI (dashed black lines). The resulting mean number of individuals contacted after correcting for reciprocity is shown in Figure 4.7. There are three main features apparent from the data. Firstly, the diagonal element where individuals in all age groups tend to mix assortatively

Table 4.2: Number of recorded contacts per participant per day by different characteristics

Category	Number of participants	Total number of contacts	Total individuals contacted	Mean number of contacts (IQ range)
Total sample	571	27,395	11,498	20.14 (12-24.75)
Gender				
Female	307	15,866	6099	19.87 (12-24)
Male	264	11,529	5,399	20.45 (12-25)
Age group (years)				
< 1	86	4,924	1,304	15.16 (10-18)
1 – 5	94	7,034	1,838	19.55 (13-23)
6 – 15	98	5,143	2,146	21.90 (13-28)
16 – 19	92	3,747	2,098	22.80 (14-29)
20 – 49	139	5,151	2,976	21.41 (13-28.75)
≥ 50	62	1,396	1,136	18.32 (10-24)
Day of keeping diary				
Weekday	402	20,067	8,004	19.91 (12-24)
Weekend	169	7,328	3,494	20.67 (13-28)
Location				
Kilifi Township	110	3,825	1,925	17.50 (11-22)
Matsangoni	134	6,745	2,961	22.10 (14-28)
Ngerenya	86	6,557	1,883	21.90 (13-26)
Roka	153	7,463	3,049	19.93 (11-24)
Tezo	88	2,805	1,680	19.09 (12-26)
Occupation				
Agriculture/Fishing	38	1,096	774	20.37 (12-24)
Business person	28	976	658	23.50 (14-30.5)
Casual labourer	30	928	620	20.67 (11-30)
Office worker	2	30	23	11.50 (7-16)
Pre-school	104	5,830	1,781	17.13 (11-21.5)
Retired	4	210	92	23 (18.5-27.5)
Student	143	6,698	3,320	23.22 (14-28)
Unemployed	180	10,224	3,399	18.88 (12-23)
Other	39	1,323	775	19.87 (8-27.5)
Missing	3	80	56	18.67 (17.5-20.25)

at relatively high rates except for a) the infants whose contacts with other infants is the lowest and b) individuals >50 years. This pattern is most pronounced in the age group 6-14 years and least pronounced in the infants. Secondly, there is high mixing recorded between infants and a) children aged 6-14 years and to a lesser degree 1-4 years old and b) adults aged 21-49 years. This pattern represents infants mixing with primary school going children (6-14 years) and adults (21-49 years) with the contacts probably happening at home since that is where the infant is likely to be.

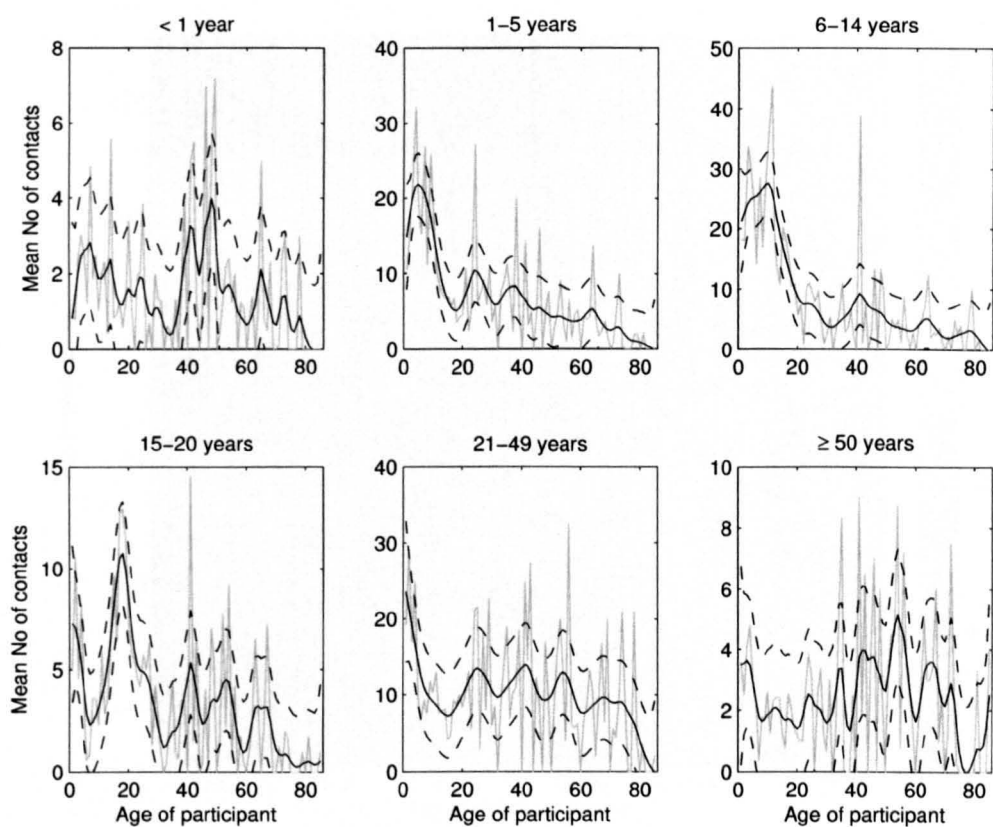


Figure 4.6: Mean number of people contacted per participant with unsmoothed contact rate (grey), smoothed (black solid lines) and the 95% confidence interval (grey dashed line). Each subplot represents the age class of the respondent.

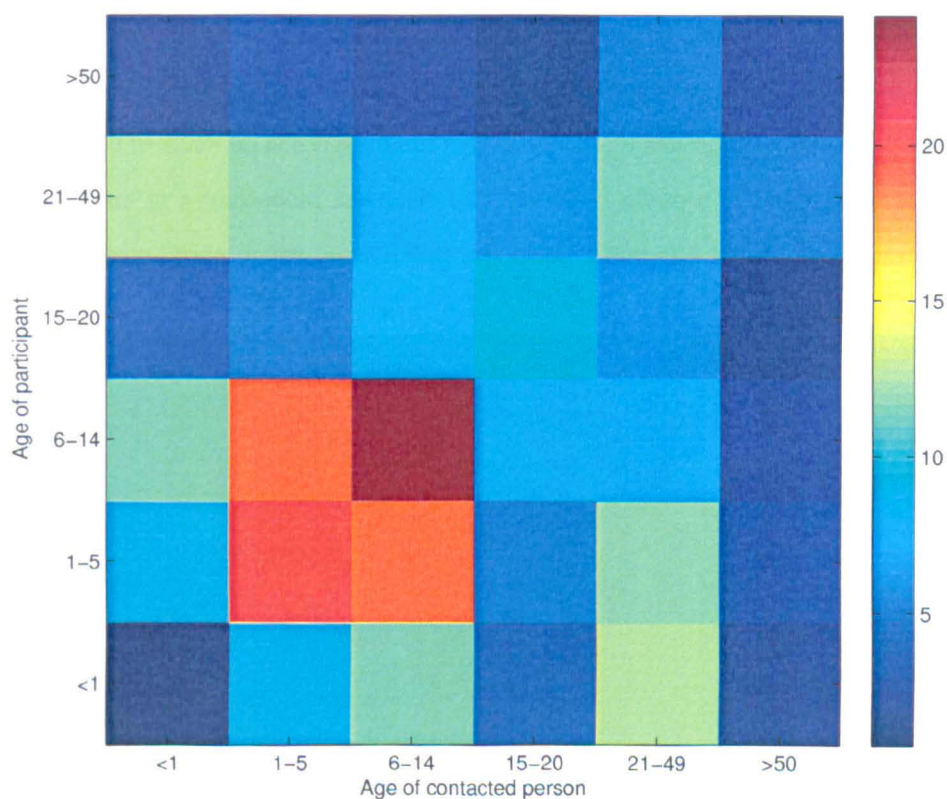


Figure 4.7: Daily contact rate with different individuals by age class accounting for reciprocity. The colourbar on the right hand side of the figure shows the mean number of contacts an individual in age class i has with respondents in age class j per day.

4.4.2 Synthetic mixing results

The synthetic WAIFW matrix is generated by combining the household, school and general population mixing matrices in a linear fashion. The WAIFW matrix resulting from this linear combination is discussed in Chapter 6. In this section, we will limit our discussion to the pattern resulting from the household mixing. Figure 4.8 shows the resulting mixing matrix from the household data for the entire KHDSS. The scale on the right hand side shows the mean number of contacts that a member in a household has scaled by the total number of individuals in the KHDSS population of the participants age class. From this figure, we can see

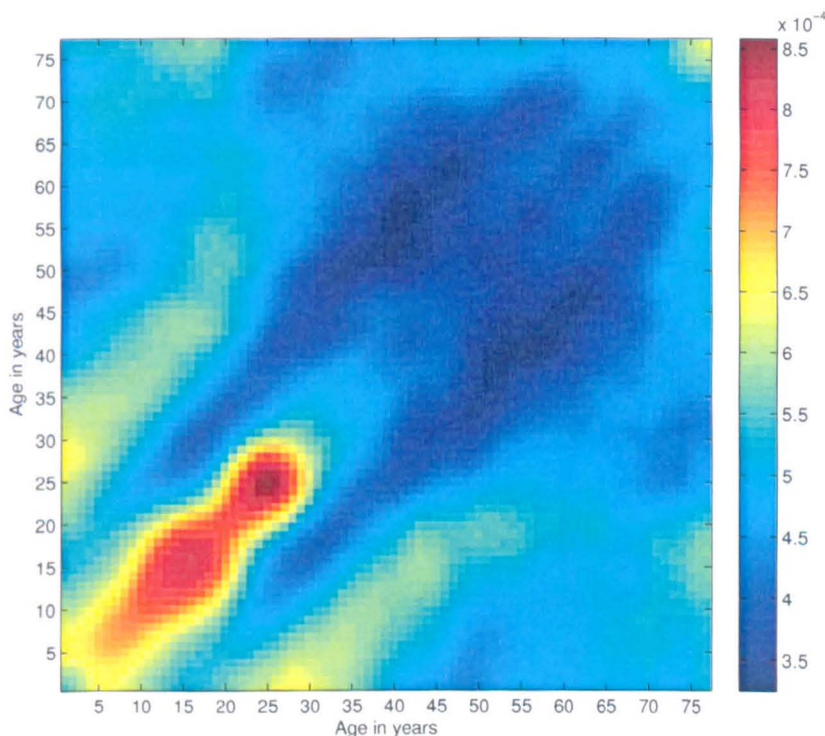


Figure 4.8: Household contact surface deduced from the Kilifi HDSS household occupancy data for the period from September 2010 to January 2011

that majority of the contacts occurring in the households within the population are among individuals aged between 5 and 26 years old. We can also observe assortative mixing between individuals of the same age class as demonstrated by the strong diagonal component up to the age of approximately 30 years. Parallel to the main diagonal on either side, there are strong mixing components between individuals under 20 years of age with individuals roughly 20-30 years old i.e. reflecting the inter-generational for of household occupancy. The peak of this mixing between younger and older individuals in between infants and approximately 30 years old individuals. These are mostly likely to be the parents or the guardians taking care of the infants in the household. Another notable feature of the mixing within the households is the strong mixing between children aged approximately 10-20 years old with elderly adults aged greater than 70 years. This result most likely reflects the tendency of grandparents living within the households of their children at old age within this population.

The diary mixing matrix has been validated against the the synthetic household mixing matrix by correlating its elements against those of the household mixing matrix. We found that they are linearly correlated with a coefficient of determination (R^2) of 0.55. Figure 4.9 shows the scatter plot from which it can be seen that a linear relationship exists between the two types of contact data. It is important to note that unlike the diary contacts matrix, the household mixing matrix is not affected by sampling error since the survey enumerates the entire population.

4.4.2.1 Contact surface by administrative location

Figure 4.8 shows the mixing pattern in the households in the entire KHDSS. To explore any differences in the household contacts, including any demographic tran-

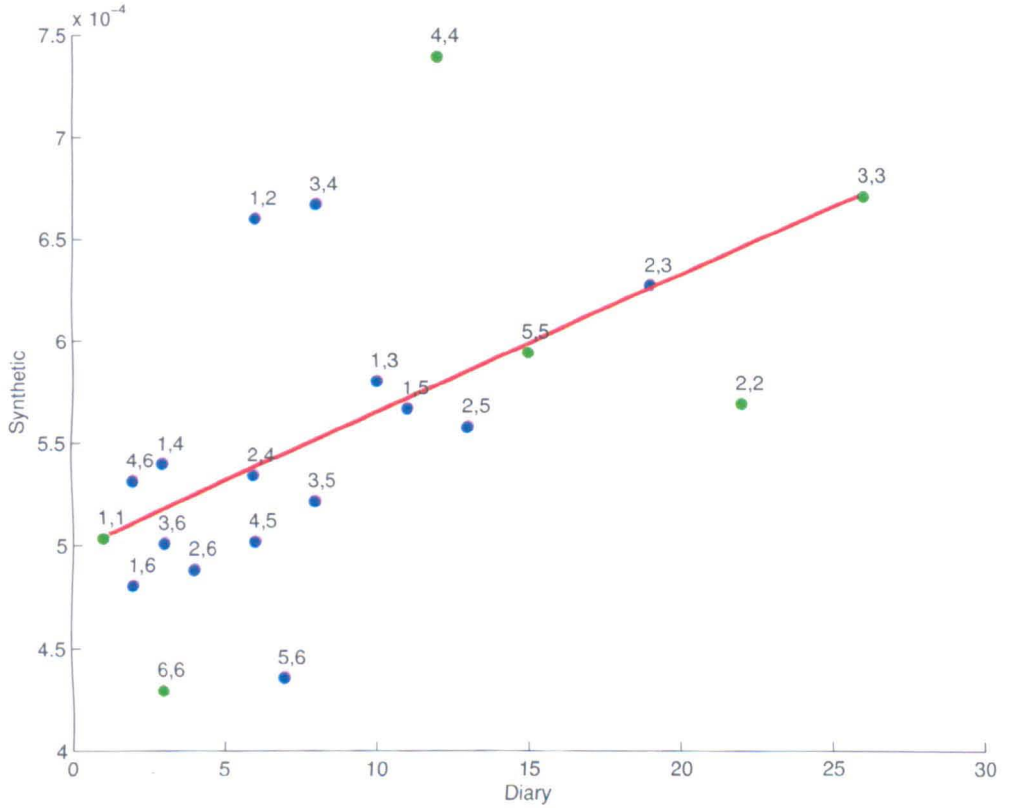
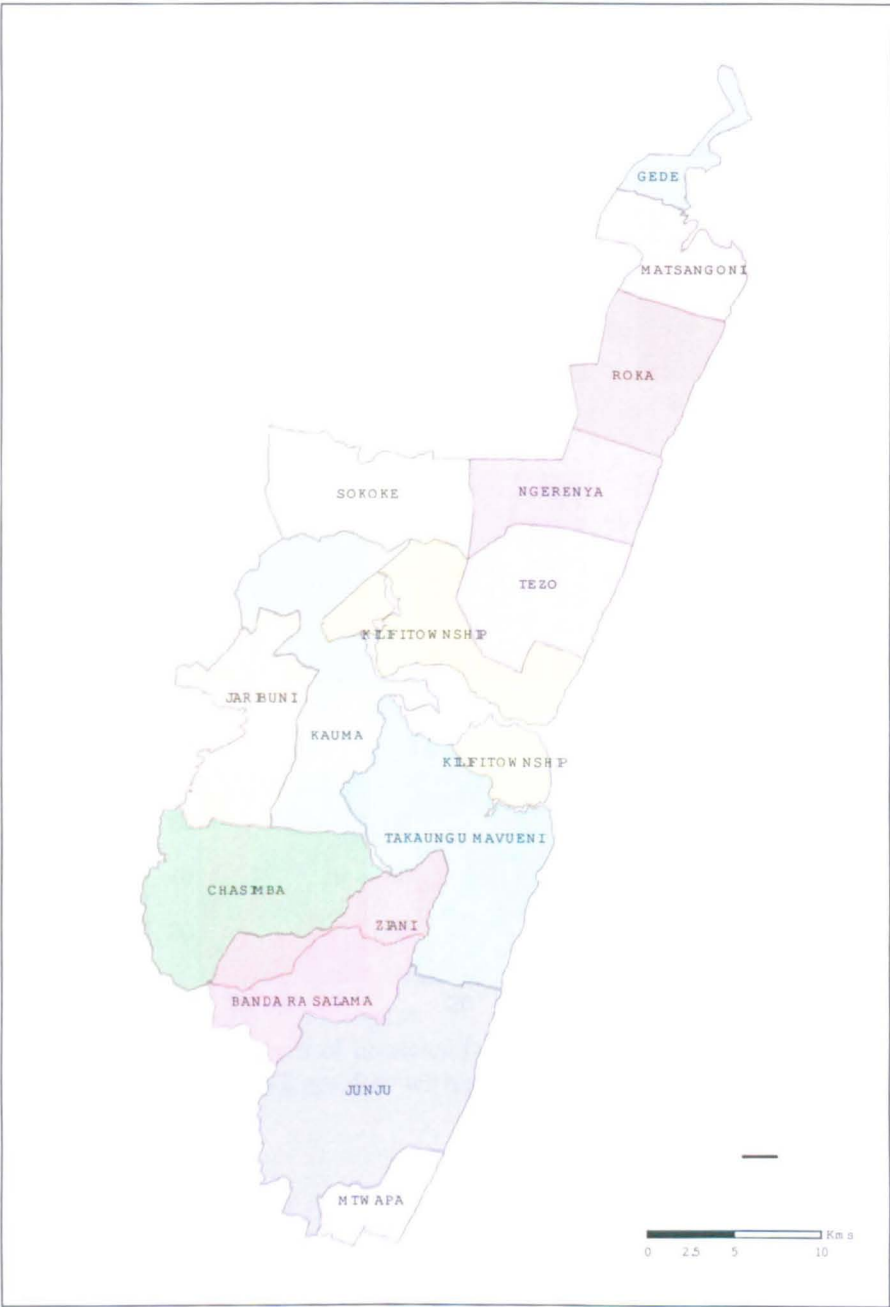


Figure 4.9: Scatter plot of diary mixing matrix elements against the household mixing matrix element with the red line showing the linear regression model fit. The green circles show the elements on the diagonal i.e. mixing between the same age classes while the blue circles represent mixing between different age groups e.g. 4,6 is the contact between individuals in secondary school (16-19 years) and elderly (> 50 years). The x-axis is the mean number of people contacted per participant per day while the y-axis is mean number of contacts that a member in a household has scaled by the total number of individuals in the KHDSS population of the participants age class

sitions, we divided the household occupancy data by the administrative locations (government administrative unit) shown in Figure 4.10. Figure 4.11 shows the contact surfaces for the different locations. To enable comparison, we used the same coloring limits for all the plots. As in the general population, we can observe assortative mixing within all the administrative locations with strong off diagonal mixing. For Kilifi Township, which is classified as an urban area, the pattern is dominated by assortative mixing among adults of age 20-30. This may be an indication of adults sharing living arrangements with fewer children compared to the rural areas. In fact, Kilifi Township has the lowest mean number of occupants per household at 6.7 while Jaribuni has the highest mean number of occupants per household at 11.26 as shown in Figure 4.12



120
Figure 4.10: Administrative locations in the KHDSS

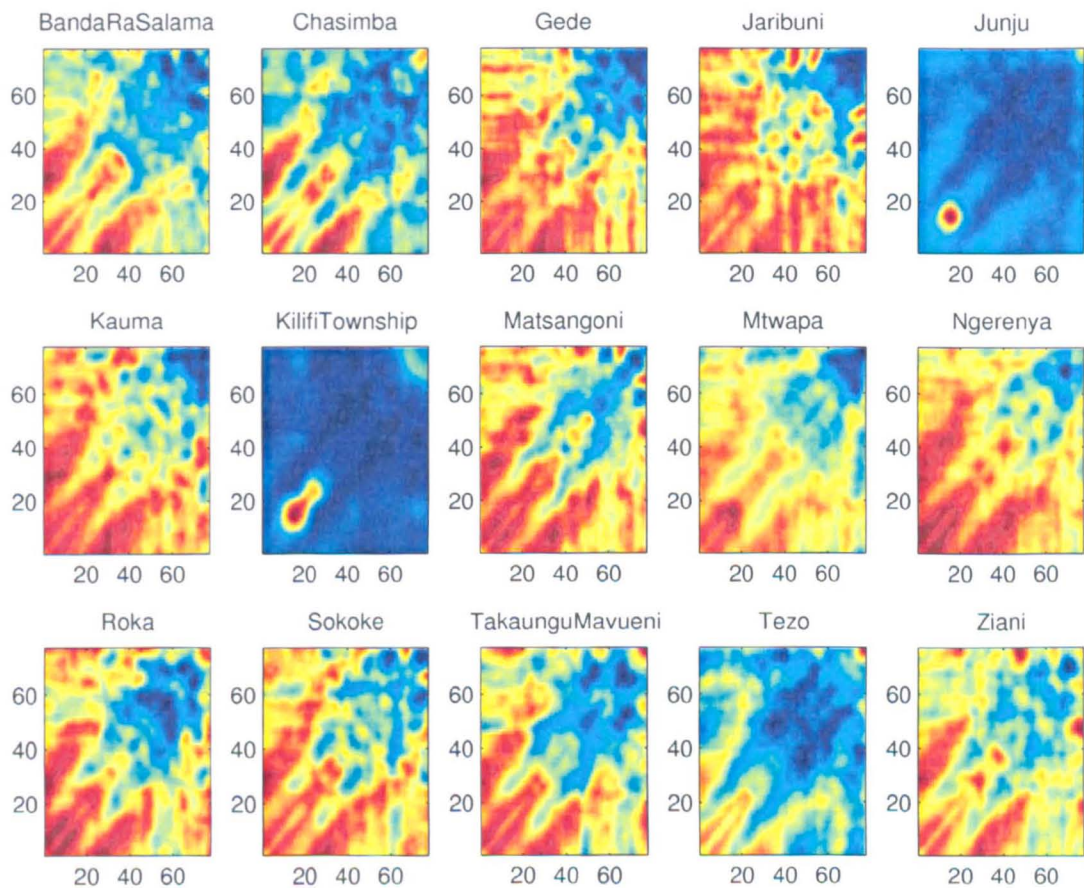


Figure 4.11: Deduced patterns of household mixing within the different administrative locations shown in Figure 4.10 within the KHDSS between Sept 2010 and Jan 2011

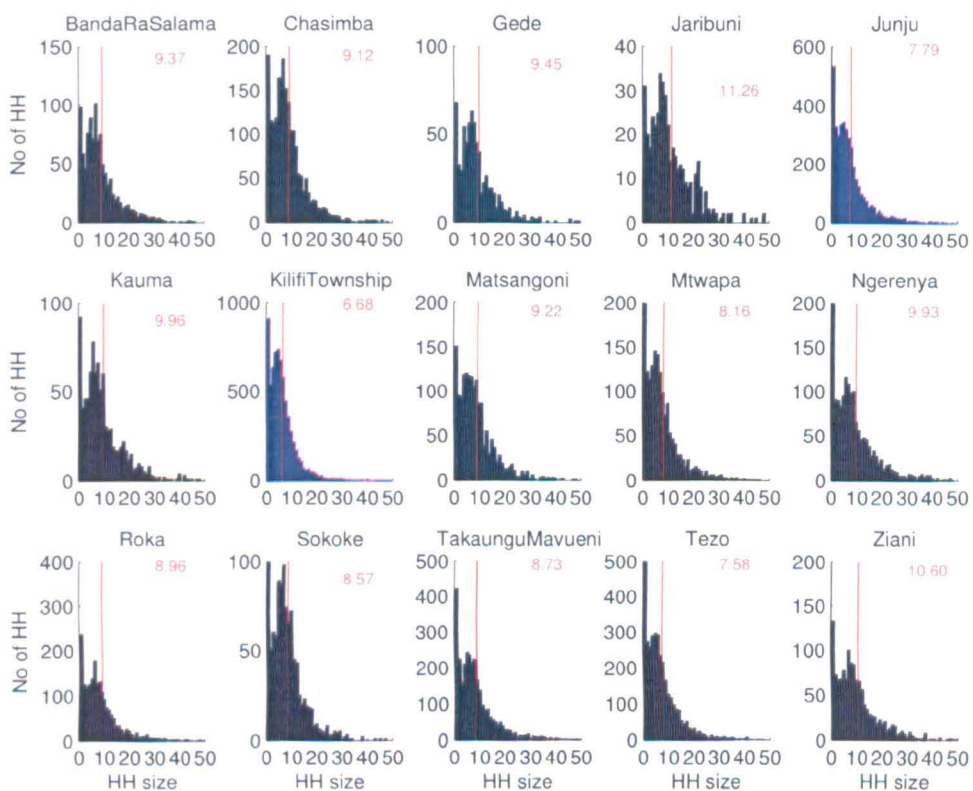


Figure 4.12: Distribution of household sizes by location. Red line and the red figure in each subplot shows the mean number of occupants in a household within the KHDSS between Sept 2010 and Jan 2011

4.5 Discussion

We report on a study designed to estimate the age specific contact rates between individuals in a defined population for the purpose of using the Who Contacts Whom (WCW) matrix to estimate the Who Acquires Infection From Whom (WAIFW) matrix. The WAIFW matrix is then used to parameterize the age-specific RSV mathematical model developed in Chapter 6. To enable the quantification of the contacts, it was necessary to define a contact event that would be easily understood and easy to record in a quantifiable way. For the purpose of the study, we defined a contact between two individuals as a direct physical contact involving some form of touch e.g. kissing, sleeping in the same bed, handshake e.t.c. These close physical contacts are the ones that are most likely to lead to a potential transmission event for RSV for which transmission has been shown to effective through close contacts [86]. What we do not include though, is the possible role of fomites which might enable relatively 'long distance' transmission possible.

The distribution of the number of total contacts were highly skewed as shown in Figure 4.4 with the majority of the participants recording more contacts of frequency 1. The distribution of the number of physical contacts with unique individuals per day is also skewed to the right with a mean of 20.14 per person per day. This value is higher than in previously published contact studies e.g. the mean number of contacts per participant per day in Vietnam [100], was 7.7 and in South Africa [110] it was 15.84. The POLYMOD study [136] also reported a mean daily contact rate (of both physical and conversational contacts) which is lower, 13.4, than in our study and the contact rate of the physical contacts alone would be expected to be much lower although not reported. There are at

least two possible reasons why this may be the case. Firstly, there may be more contacts reported within the KHDSS population due to the tradition of extended family members staying within the same compound and sharing a kitchen. This would expose them to more contacts within the household compared to people in living arrangements where only the nuclear family lives together. Secondly, the study design could have been an influence in that people were reminded to fill in the diaries at hourly intervals using programmed watches. Both of these reasons may have improved participant response rates to filling the diaries. The mean number of contacts per participant was associated with the participant's age class with the lowest mean number of contacts recorded by infants and the highest recorded by primary school (21.9) followed by the secondary school (22.8) children. Additionally, it was observed that the contacts made with pre-school and primary aged children are more assortative compared to contacts made by other age groups. This means that most of the contacts that individuals in ages 1-14 make are with people of the same age group and this result has been observed in the POLYMOD study done in several European countries [136]. This outcome is possibly the reason why young children have been shown to be important in the initial spread of respiratory infections requiring close contacts [196].

One of the main assumptions in our work is that physical contact with another person is what has been defined as the main at-risk event for transmission of an infection. However, there may exist other at-risk events that this methodology has not captured e.g. the duration and intimacy of a contact and being in a confined space with another person but not touching them [16, 167] and the possible transmission through fomites via the sharing of contaminated items [83]. Such at-risk events are probably less important in the transmission of RSV compared to

the physical contacts captured by our work. Another limitation of diary study is that approximately 60% of the participants had their diaries filled in by a shadow. Shadowing has the disadvantage that it can lead to behaviour modification of the person being shadowed. Additionally, a shadow would not necessarily record all contacts since they are not always in the company of the participant. Although every effort was made to minimize recall bias, it is possible that the data is subject to it. Participants were required to fill in the diary as often as they could during the course of the day and they had wrist watches that had a timed alarm that went off at a certain interval to serve as a reminder. The simplicity of our study in terms of the small number of actions that the participants had to do is likely to be an advantage relative to previous studies requiring the participants to record e.g. location of contact, type of contact, duration. This ensured that the participants effort was geared towards only a few entries i.e. the age of the respondent and the frequency of the contacts and it is possible that it enhanced the correctness of the data entered. A limitation that was reported in one of the main diary contact studies (POLYMOD study) is that of right censoring. In that study, data analyzed were right censored at 29 contacts because of a limited number of possible diary entries in some of the study populations. However, in our study, this effect of right censoring is minimized if not absent. This is because there are only three individuals who reported 68 contacts with different people while the diary had a capacity to record a total of 68 different contacts. Participants in the study were provided with a reminder table on which they would record any additional contacts incase the diary spaces were filled up. However, during the exit interview, it was established that none of the three had recorded any contacts in the reminder table that were not in the main diary. None of the participants expressed any concerns

about limited space in the diaries.

Another limitation that is also worth highlighting is that of a high non-consent rate. We reported a non-consent rate of 50.1%. Non-consent rate was determined as the proportion of all cases in which a respondent refuses to consent and cuts off contact with the field worker well before keeping the diary [61]. It is important to note, as shown in Table 4.1 that majority of the reasons for non-consenting in the study are not known, followed by individuals who are temporarily away and then individuals who are generally not interested in the study. Although there is a high non-consent rate, there are actually multiple reasons and only in part due to refusals. It is important to note that since there was a fixed time schedule for sampling, there could be very little to carry over from one month of the study to the next and this would increase non-participation. More important in the assessment of the influence of non-participation bias is the extent to which non-participation is associated with exposure to varying contact rate. We did not record any data from the individuals who did not consent and therefore it would be difficult to determine if there are any differences in the contact patterns between the participants and those who did not consent to participate. However, a comparison between the synthetic and the diary data suggests that the pattern of contacts is similar for the two methods and since the synthetic method enumerates the entire population, then it suggests the individuals who did not consent may have very likely had the same mixing pattern as the participants.

Using contacts diaries in the population has been shown to be a feasible method of collecting social contact data. However, we have also used household occupancy data to try and validate our finding using a different approach. From Figure 4.8, the synthetic household mixing represents two clear features: 1) a dominant

diagonal from 5 year old children to 26 year old individuals and 2) there is an upper and lower off diagonals possibly accounting for contacts that parents have with children. This patterns has also been observed by Fumanelli et al [60] with data from European countries. This mixing patterns also seems to be conserved across the different administrative locations within the KHDSS as seen in Figure 4.11 with the exception of Kilifi Township and Junju. The pattern in the two locations is indicative of more adults sharing living arrangements compared to the other locations within the Kilifi HDSS. This social contact matrix from the diary survey has been validated against this household synthetic matrix by jointly regressing their elements. Although there is a good agreement between the two matrices, their elements differ by a factor i.e. the diary matrix records more contacts compared to the household mixing matrix though the mixing pattern remains the same. This is expected since the synthetic household matrix is composed of only the contacts in the household and ignores all the contacts occurring elsewhere e.g. between households. Given that the pattern observed in the two matrices is closely related, it possibly implies two things either the diary contacts are dominated by contacts within the household or contacts outside of the household are assortative in nature. From this work however, it is not possible to discriminate between the two suggestions. The use of the synthetic matrix for the determination of social contacts relevant for the transmission of close contacts has the advantage that it is cheap since it makes use of available social demographic data. This method can be extended by using data from other sources e.g. school and work attendance data as proposed by [60] and therefore form a good alternative for settings without social contact data. Historic mixing patterns can also be constructed and used, together with mathematical models of disease transmission, to study the effect of

demographic transition on the transmission of infections.

Chapter 5

A mathematical model of RSV transmission dynamics

5.1 Introduction

Mathematical models in the description of biological processes have previously been developed for two reasons. The first one is a predictive purpose [180, 12, 130] where the model is developed to include sufficient processes parameterized from observational and experimental data with the aim of predicting the effect of an action or intervention. The second purpose for which a model would be developed is to gain an understanding of the underlying process influencing the behaviour of the system [69, 46, 131]. The complexity of such models is kept to a minimum and at the same time ensuring that the most important aspects of the infection are captured [72]. The model presented in this chapter is one of the second nature and also serves as a template on to which further biological and demographical complexity can be added. A simple mathematical model is presented which allows

for the definition of the most important epidemiological elements of RSV. The compartmental model developed allows the host population to be categorized based on their infection history and immunity status.

There have been previous RSV mathematical models developed. Weber and others [201] compared a standard SIR model with a more realistic model of RSV transmission in which individuals acquire immunity gradually after repeated exposure. The risk of an exposed person contracting an infection was assumed to decrease after the first four experiences such that the risk of infection is 50% for the second, 35% for the third and 25% for the fourth infection compared to primary infection. Qualitatively, the two models gave an equally good fit to a time series of hospital case reports. Estimates of the basic reproduction number (R_0) ranged from 1.2 to 2.1 with the SIRS model and 5.4 to 7.1 with the model accounting for gradual acquisition of partial immunity.

Luis and others [1] developed a mathematical model of RSV with two age classes describing the transmission of RSV in Valencia, Spain. They fitted the model to hospitalization data from illnesses related to RSV from Valencia. They further considered a new born vaccination strategy in terms of the estimated cost of vaccine, the average cost of hospitalization of RSV infected children who develop acute symptoms and the parent work loss. From the outcome of the model, a reduction of 2 million Euro of total cost is predicted or an estimation of 3 days of parent work loss on average for hospitalized infected children. The predictive results arising from this work should be looked at in the light of the following limitations. Firstly by using a SIRS model, they have assumed that the primary and secondary infections are similar in terms of their recovery period and infectiousness. This is not necessarily true as it has been shown that the secondary

infectious period can be reduced by as much as 40% from the primary infectious period [158, 156]. The assumption will potentially bias the simulation and vaccine effectiveness results. For example, the number of cases hospitalized might be overestimated since it will include both primary and secondary infecteds in the first age cohort. Secondly, the vaccination strategy is evaluated at birth yet we know from previous work that the efficacy of a vaccine given at this age may be compromised for possibly two reasons. The first one is that the immune system of the infant may not be properly developed to elicit an immune response [174] and the second is that the presence of RSV specific maternally derived antibodies may influence the effect of the vaccine given at that age hence leading to possible vaccine failures [187, 160].

White and others [205] proposed a single model structure that captures four possible host responses: partial susceptibility, altered infection duration, reduced infectiousness and temporary immunity to infection. By setting the homotopic parameters to extreme values, the model generated a set of eight nested sub-models. These models were applied to time series case reports from eight geographically distinct locations. Models that incorporated either of the two extreme assumptions of immunity (none or solid and lifelong) were unable to reproduce the observed temporal dynamics. Models with waning or partial immunity to disease were both visually comparable with the best fitting model being the one with lifelong partial immunity to RSV infection. This work therefore seems to suggest that the data supports two model forms: a) lifelong partial immunity and b) waning immunity.

The model that we present in this chapter will seek to address some of the limitations of the previously developed RSV models. We have considered the natural progression of the infection from previous studies and have endeavored

to keep the structure of the model as simple as possible to allow for analytical tractability but not so simple as to exclude the most important epidemiological aspects. We have presented the model in two forms in this chapter. The first one is an analytic evaluation to determine the stability of the model and the conditions for the invasion of the infection in an infection free population. Given the analytical intractability of the endemic equilibrium model, the second form of the model is presented in terms of numerical simulations exploring the behaviour of the model in the presence of the infection as well as to confirm the analytical results both in the presence and the absence of a vaccination programme.

5.1.1 Objectives

The objectives of the work presented in this chapter are:

- to model the transmission dynamics of RSV using a compartmental mathematical model and explore the stability behaviour both in the disease free population and in the infection endemic state.
- to explore the behaviour of the model under different parameter values and explore regions exhibiting backward and forward bifurcations. Additionally, I have considered whether the criteria for multiple endemic equilibria is satisfied within a set of biologically meaningful parameters.

5.2 Model structure

I have developed a deterministic compartmental model in which an individual occupies one and only one compartment. This kind of compartmental deterministic

modelling has previously been used to describe childhood diseases [180, 164, 125]. The host population is divided into nine distinct epidemiological groups according to their infection status. The cartoon in Figure 5.1 shows the flow of individuals between the different compartments with respect to time.

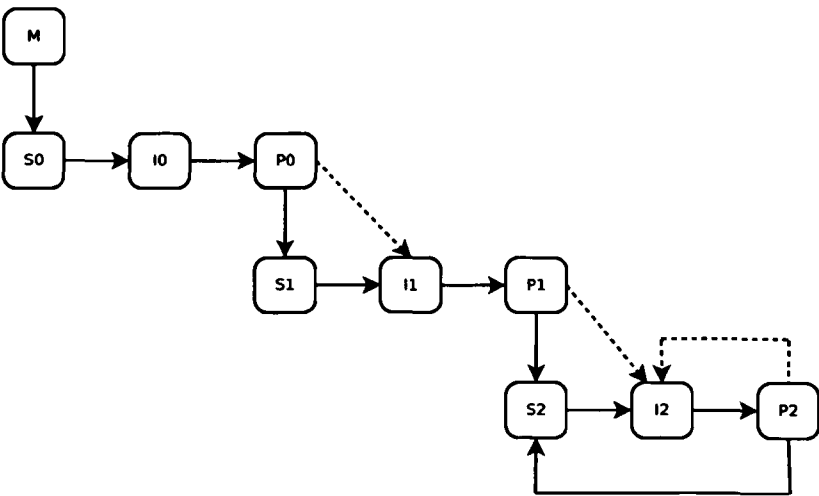


Figure 5.1: Schematic flow diagram of the compartmental model. Arrows represent the flow of individuals between states. The parameters defining the rates of flow are discussed in the text and listed in Table 5.1.

Simply, the infection dynamics are as follows: Individuals are born into compartment M during which they are protected from contracting the infection by maternal antibodies. It has previously been shown that children born to seropositive mothers are born with maternally derived antibodies against RSV [154, 152, 20]. In one prospective study [154], newborn infants were examined prospectively for one year for evidence of infection with RSV. Mean titre of maternal IgG antibody to RSV was significantly higher in those mothers whose babies remained uninfected. Babies born to mothers with high levels of IgG antibody to RSV were protected against infection with the virus during the first months of life when the

risk of severe disease was highest. In another study measuring the duration of RSV specific maternal IgG in infants [152], 97% of the children in the birth cohort had detectable levels of maternal antibodies at birth.

The maternal antibodies wane leaving individuals susceptible to primary infection, S_0 , upon which they progress into primary infected class, I_0 , at a rate λ as shown in Eqn.(5.1). Individuals then recover into class P_0 at a rate γ_0 where they have solid but waning immunity. Individuals are known to be repeatedly infected throughout life [68, 87, 149] and so immunity is not solid and/or life-long. Individuals then lose their temporary protection to become susceptible to a second infection at a reduced rate $\sigma_0\lambda$. Re-infected individuals then progress to secondary infected class I_1 with reduced infectiousness compared to I_0 possibly due to reduced duration of shedding and reduced viral shedding. A recent study [158] investigating the duration of RSV infection and viral shedding in relation to the infection history, age and severity showed that the rate of recovery was 40% faster for children previously infected. Transmission has also been shown to be associated with the quantity of viral material shed [85]. Once individuals recover from the second infection, they enter class P_1 where they lose their protection at a rate ρ_1 to become susceptible to a re-infection and proceed to class I_2 . The rate of infection from S_2 is $\sigma_1\lambda$ such that $\sigma_1\lambda \leq \sigma_0\lambda \leq \lambda$. This implies that there is long-term partial immunity conferred upon an individual by a previous infection. The risk of an exposed person getting infected decreases after the first and second experiences of an infection and subsequent re-infections are assumed not to confer any extra protection and hence individuals will move between compartments S_2 , I_2 and P_2 . This assumption is based on a longitudinal study done on a birth cohort during seven successive epidemics where the attack rate for the second and third

infections was reduced by 25% and 35% respectively, with statistical significance, compared to the attack rate of the primary susceptible individuals [93]. There was no extra benefit acquired beyond the third infection and the reduction in the rate of attack was minimal and without statistical significance.

Eqn.(5.1) shows the set of Ordinary Differential Equation (ODE) representing the flow of individuals through the 10 epidemiological classes as represented in Figure 5.1. The model parameters, their description and values are shown in Table 5.1. The model parameters in the table are not strictly based on the review done in Chapter 3. This is because they were chosen so as to give a basic reproduction number of approximately 7.3 which is within what previous modelling work suggested [205, 201]. However, the parameters used in Chapter 6 and the upper and lower bounds used in the uncertainty and sensitivity analysis have been based on the review. During the model numerical simulations, the population size was assumed constant. The mortality and fertility rates are therefore assumed constant and equal, denoted by μ , such that the value is the reciprocal of the average life expectancy. The total number of deaths is given as μN and the number of births is chosen so as to match the mortality rate. The initial conditions are such that 99% of the people are in the M class and 1% are in the I_0 class.

Table 5.1: Baseline parameter estimates used in the numerical simulations in Figure 5.2

Symbol	Description	Value	REF
μ	Birth/death rate	Estimated as the reciprocal of the life expectancy which is assumed to be 50 years	$\mu = 0.02$ World Bank Data
ω	Rate of decay of maternal antibodies	$\omega = 4/yr$. The duration is approximately 3 months	$\omega = 4/yr$ [78]
b	Transmission parameter	$\lambda = b \left[\frac{\alpha_0 I_0 + \alpha_1 I_1 + \alpha_2 I_2}{N} \right]$ where N is the population size and λ is the force of infection	$b = 300$ Taken to give an $R_0=7.36$
σ_i	Long-term immunity factor (partial) reducing the susceptibility of previously exposed individuals in S_i where $i=2,3$ and P_2 classes	$0 \leq \sigma_0, \sigma_1, \sigma_2 \leq 1$ and $\sigma_2, \sigma_1 \leq \sigma_0$. $\sigma_i = 1$ implies complete immunity $\sigma_i = 0$ implies no immunity where $i=0,1,2$	$\sigma_0 = 1$ $\sigma_1 = 0.75$ $\sigma_2 = 0.65$ [93]
ρ_i	Rate of waning of short term immunity of recovered individuals, ρ_i where $i=0,1,2$	To be explored. Assume protection for 3 months and 6 months after primary infection	$\rho_0 = 4/yr$ $\rho_1 = 2/yr$ $\rho_2 = 2/yr$ [184]
η_i	Short-term immunity factor reducing the susceptibility of recovered individuals in the P_i classes	$0 \leq \eta_i \leq 1$ $\eta_i = 0$ implies complete immunity $\eta_i = 1$ implies no additional immunity. $\eta_2 \leq \eta_1 \leq \eta_0$	$\eta_0 = 0.5$ $\eta_1 = 0.25$ $\eta_2 = 0.25$ Assumption
γ_i	Rate of recovery from infected classes I_i into recovered classes P_i	$\gamma_1, \gamma_2 \leq \gamma_0$ Assume 9 days and then approx 40% reduction to 4.2 days	$\gamma_0 = 40.6/yr$ $\gamma_1 = 86.9/yr$ $\gamma_2 = 86.9/yr$ [199, 85] [158]
α_i	Factor reducing infectiousness	$\alpha_0 = 1$ and $0 \leq \alpha_1, \alpha_2 \leq 1$ $\alpha_i = 0$ implies not infectious and $\alpha_i = 1$ implies complete infectiousness	$\alpha_0 = 1$ $\alpha_1 = 0.5$ $\alpha_2 = 0.5$ Assumption based on [158]

$$\begin{aligned}
\frac{dM}{dt} &= \mu N - M(\mu + \omega) \\
\frac{dS_0}{dt} &= \omega M - S_0(\mu + \lambda) \\
\frac{dI_0}{dt} &= \lambda S_0 - I_0(\mu + \gamma_0) \\
\frac{dP_0}{dt} &= \gamma_0 I_0 - P_0(\mu + \rho_0 + \eta_0 \sigma_0 \lambda) \\
\frac{dS_1}{dt} &= \rho_0 P_0 - S_1(\mu + \sigma_0 \lambda) \\
\frac{dI_1}{dt} &= \sigma_0 \lambda S_1 + \eta_0 \sigma_0 \lambda P_0 - I_1(\mu + \gamma_1) \\
\frac{dP_1}{dt} &= \gamma_1 I_1 - P_1(\mu + \rho_1 + \eta_1 \sigma_1 \lambda) \\
\frac{dS_2}{dt} &= \rho_1 P_1 + \rho_2 P_2 - S_2(\mu + \sigma_1 \lambda) \\
\frac{dI_2}{dt} &= \eta_1 \sigma_1 \lambda P_1 + \sigma_1 \lambda S_1 + \eta_2 \sigma_2 \lambda P_2 - I_2(\mu + \gamma_2) \\
\frac{dP_2}{dt} &= \gamma_2 I_2 - P_2(\mu + \rho_2 + \eta_2 \sigma_2 \lambda)
\end{aligned}$$

$$\text{where} \quad \lambda = \frac{b}{N} \sum_{i=0}^2 \alpha_i I_i \tag{5.1}$$

In the next section, I will present an analytical analysis of the model exploring the basic reproduction number, the invasion threshold and the conditions under which multiple sub or supercritical endemic equilibria can exist. A numerical solution to the Eqn.(5.1) is also presented and the results compared with the analytical results.

5.3 Results

5.3.1 Invasion threshold

The basic reproduction number is a well known measure representing the potential for an infection to be transmitted in a population. Technically, we denote it as R_0 in this thesis and define it as the average number of new infections that arise from an average primary case during their infectious period in a completely susceptible population [116, 10]. In order to calculate the analytical invasion threshold, we need to work out the closed form of R_0 and then equate it to one. The R_0 has been calculated based on the standard methodology illustrated in [42, 195].

Let's consider a homogeneous population whose individuals can be distinguished by their infection status as shown in Figure 5.1. The compartments are mutually disjoint and an individual can belong to one and only one compartment at any unit time. The basic reproduction number cannot be determined from the structure of the model alone but also depends on the definition of infected and infectious compartments. Therefore, the model has been defined starting with the three infected classes and then the disease free compartments such that the solution can be expressed as $[I_0^*, I_1^*, I_2^*, M^*, S_0^*, P_0^*, S_1^*, P_1^*, S_2^*, P_2^*]$.

In order to compute R_0 , it is important to distinguish new infections from all other changes in the population. To do this, let:

$G_i(x)$ be the rates of appearance of new infections in compartment i

$V_i^+(x)$ be the rate of transfer of individuals into compartment i by all other means

$V_i^-(x)$ be the rate of transfer of individuals out of compartment i and $V_i = V_i^- - V_i^+$

Therefore the set of ODEs in Eqn.(5.1) can generally be written as:

$\dot{X}_i = G_i(x) - V_i(x)$ where $i = 1, 2, \dots, 10$. From Eqn.(5.1), then we generated the

following vectors describing the flow of individuals between compartments:

$$G(x) = \begin{bmatrix} \lambda S_0 \\ \sigma_0 \lambda S_1 + \eta_0 \sigma_0 \lambda P_0 \\ \eta_1 \sigma_1 \lambda P_1 + \sigma_1 \lambda S_2 + \eta_2 \sigma_2 \lambda P_2 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, V(x) = \begin{bmatrix} I_0(\mu + \gamma_0) \\ I_1(\mu + \gamma_1) \\ I_2(\mu + \gamma_2) \\ -\mu N + M(\mu + \omega) \\ -\omega M + S_0(\mu + \lambda) \\ -\gamma_0 I_0 + P_0(\mu + \rho_0 + \eta_0 \sigma_0 \lambda) \\ -\rho_0 P_0 + S_1(\mu + \sigma_0 \lambda) \\ -\gamma_1 I_1 + P_1(\mu + \rho_1 + \eta_1 \sigma_1 \lambda) \\ -\rho_1 P_1 - \rho_2 P_2 + S_2(\mu + \sigma_1 \lambda) \\ -\gamma_2 I_2 + P_2(\mu + \rho_2 + \eta_2 \sigma_2 \lambda) \end{bmatrix}$$

The infected compartments are I_0 , I_1 and I_2 . Assuming that there is no maternally protected class, then the form of the disease free equilibrium solution, $X_0 = [0, 0, 0, S_0^*, 0, 0, 0, 0, 0]$, will give us an R_0 of a simple SIR model: $R_0 = \frac{\alpha_0 b}{\gamma_0 + \mu}$. The disease free equilibrium (DFE) solution, such that $I_0 = I_1 = I_2 = 0$, in the presence of a maternally protected class has the form $X_0 = [0, 0, 0, M^*, S_0^*, 0, 0, 0, 0]$. It is worth indicating at this point that maternal protection at the DFE equilibrium does not make biological sense since in the absence of infection, there would be no RSV specific maternal antibodies against the infection. However, the assumptions allows us to calculate the R_0 of an infection. Without loss of generality, assume we scale the population to unity. We then work out the disease free stable equilibrium point of our model by equating to zero the differential equations describing the maternally protected and the primary susceptible classes and solving the resulting

simultaneous equations. Doing so gives us the following stable equilibrium point $X_0 = [0, 0, 0, \frac{\mu}{\mu+\omega}, \frac{\omega}{\mu+\omega}, 0, 0, 0, 0]$. From [42, 74], the basic reproduction number is calculated as $R_0 = \vartheta(FV^{-1})$ where ϑ is the spectral radius and $F = \left[\frac{\partial G_i}{\partial x_j}(x_0) \right]$ and $V^{-1} = \left[\frac{\partial V_i}{\partial x_j}(x_0) \right]$ are the Jacobi evaluated at the disease free equilibrium. The spectral radius of a square matrix is defined as the supremum of the elements in its set of eigenvalues. The spectral radius will therefore be the dominant eigenvalue of the FV^{-1} matrix. So F and V^{-1} become:

$$F = \begin{bmatrix} \frac{\alpha_0 b \omega}{\mu + \omega} & \frac{\alpha_1 b \omega}{\mu + \omega} & \frac{\alpha_2 b \omega}{\mu + \omega} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, V^{-1} = \begin{bmatrix} \frac{1}{\gamma_0 + \mu} & 0 & 0 \\ 0 & \frac{1}{\gamma_1 + \mu} & 0 \\ 0 & 0 & \frac{1}{\gamma_2 + \mu} \end{bmatrix}$$

Consequently, we evaluate the product FV^{-1} . So, FV^{-1} becomes:

$$FV^{-1} = \begin{bmatrix} \frac{\alpha_0 b \omega}{(\gamma_0 + \mu)(\mu + \omega)} & \frac{\alpha_1 b \omega}{(\gamma_1 + \mu)(\mu + \omega)} & \frac{\alpha_2 b \omega}{(\gamma_2 + \mu)(\mu + \omega)} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5.2)$$

To interpret the entries of matrix (5.2), lets consider an infected individual introduced in a disease free population in compartment k . The (j, k) entry in V^{-1} is the average length of time this individual spends in compartment j . The (i, j) entry of the F matrix is the rate at which infected individuals in compartment j produce new infections in compartment i . The (i, k) entry of matrix (5.2) is the expected number of new infections in compartment i produced by the infected individual originally introduced into compartment k and the matrix is referred to as the next generation matrix as defined in [195]. To calculate R_0 , I evaluate the spectral radius which is the dominant eigen value of matrix (5.2). To calculate the closed

form of the eigen values, I used the symbolic math toolbox in MATLAB® [128] powered by the MuPAD symbolic engine. The resulting three eigen values of FV^{-1} are denoted as χ_i where $i = 1, 2, 3$. $\chi_1 = \chi_2 = 0$ and $\chi_3 = \frac{b\alpha_0\omega}{(\mu+\gamma_0)(\mu+\omega)}$ provided that $(\mu + \gamma_0)(\mu + \omega) \neq 0$. Given that μ , ω and γ_0 are all rates greater than 0, then $(\mu + \gamma_0)(\mu + \omega)$ is always greater than 0. Again, since all the rates are greater than zero, then the $\max(|\chi_i|) = \max(\chi_i) = \chi_3$. Hence, the basic reproduction number is given as:

$$R_0 = \frac{b\alpha_0\omega}{(\mu + \gamma_0)(\mu + \omega)} \quad (5.3)$$

At the invasion threshold, $R_0 = 1$. In order to calculate the analytic invasion threshold, we equate Eqn.(5.3) to 1 and then make b , the transmission parameter, the subject of the resulting expression. So, the invasion threshold denoted as b^* , can be expressed as shown in Eqn.(5.4):

$$b^* = \frac{(\mu + \gamma_0)(\mu + \omega)}{\alpha_0\omega} \quad (5.4)$$

We will compare the results obtained from Eqn.(5.4) with numerical simulations.

5.3.2 Local stability of the model

In section 5.3.1, we have calculated the invasion threshold from the disease free equilibrium point of the model. In this section, we aim to determine the neighbourhood stability of this disease free equilibrium and the endemic state of the infection. According to [2], local or neighbourhood stability is the investigation of the behaviour of the system at equilibrium to small perturbations.

The steady states are obtained by setting Eqn.(5.1) to zero and solving the

resulting simultaneous equations. Since we are calculating the steady state in the presence of the infection in the population, then the form of the solution will be $X^* = [M^*, S_0^*, I_0^*, P_0^*, 0, 0, 0, 0, 0]$. This assumes that infection is present only in the primary infected class and this assumption allows the model to be analytically tractable even though it is biologically implausible. The Jacobian is then calculated from the first four equations in Eqn.(5.1). The Jacobian is expressed as follows:

$$\begin{bmatrix} \frac{\partial F}{\partial M}(X^*) & \frac{\partial F}{\partial S_0}(X^*) & \frac{\partial F}{\partial I_0}(X^*) & \frac{\partial F}{\partial P_0}(X^*) \\ \frac{\partial G}{\partial M}(X^*) & \frac{\partial G}{\partial S_0}(X^*) & \frac{\partial G}{\partial I_0}(X^*) & \frac{\partial G}{\partial P_0}(X^*) \\ \frac{\partial H}{\partial M}(X^*) & \frac{\partial H}{\partial S_0}(X^*) & \frac{\partial H}{\partial I_0}(X^*) & \frac{\partial H}{\partial P_0}(X^*) \\ \frac{\partial Z}{\partial M}(X^*) & \frac{\partial Z}{\partial S_0}(X^*) & \frac{\partial Z}{\partial I_0}(X^*) & \frac{\partial Z}{\partial P_0}(X^*) \end{bmatrix} \quad (5.5)$$

such that $F = \frac{dM}{dt}$, $G = \frac{dS_0}{dt}$, $H = \frac{dI_0}{dt}$ and $Z = \frac{dP_0}{dt}$. After calculating the above partial derivatives on the equation describing the rate of change of maternally protected individuals, the primary susceptible, primary infected and recovered classes we obtain the following Jacobian matrix:

$$J = \begin{bmatrix} -(\mu + \omega) & 0 & 0 & 0 \\ \omega & -\mu - \frac{b\alpha_0 I_0}{N} & -\frac{b\alpha_0 S_0}{N} & 0 \\ 0 & \frac{b\alpha_0 I_0}{N} & -\mu - \gamma_0 + \frac{b\alpha_0 S_0}{N} & 0 \\ 0 & 0 & \gamma_0 & -\mu \end{bmatrix} \quad (5.6)$$

Solving for X^* gives the following solution

$$\begin{bmatrix} M^* \\ S_0^* \\ I_0^* \\ P_0^* \end{bmatrix} = \begin{bmatrix} \frac{N\mu}{\mu+\omega} \\ \frac{N\omega}{\mu+\omega} \\ 0 \\ 0 \end{bmatrix} \text{ or } \begin{bmatrix} \frac{N\mu}{\mu+\omega} \\ \frac{N(\gamma_0+\mu)}{\alpha_0 b} \\ \frac{N\mu(\alpha_0\omega b - \gamma_0\omega - \mu\omega - \mu^2 - \gamma_0\mu)}{\alpha_0 b(\mu+\omega)(\gamma_0+\mu)} \\ \frac{N\gamma_0\omega}{(\gamma_0+\mu)(\mu+\omega)} - \frac{N\gamma_0}{\alpha_0 b} \end{bmatrix} \quad (5.7)$$

The first solution gives the disease free equilibrium. Substituting the disease free equilibrium into matrix (5.6) gives

$$J_{df} = \begin{bmatrix} -\mu - \omega & 0 & 0 & 0 \\ \omega & -\mu & \frac{-\alpha_0 b \omega}{\mu+\omega} & 0 \\ 0 & 0 & \frac{\alpha_0 b \omega}{\mu+\omega} - \mu - \gamma_0 & 0 \\ 0 & 0 & \gamma_0 & -\mu \end{bmatrix} \quad (5.8)$$

while substituting the second solution vector gives

$$J_d = \begin{bmatrix} -\mu - \omega & 0 & 0 & 0 \\ \omega & -\frac{\alpha_0 b \mu \omega}{(\mu+\omega)(\gamma_0+\mu)} & -\mu - \gamma_0 & 0 \\ 0 & \frac{\alpha_0 b \omega \mu}{(\mu+\omega)(\gamma_0+\mu)} - \mu & 0 & 0 \\ 0 & 0 & \gamma_0 & -\mu \end{bmatrix} \quad (5.9)$$

The stability of the system is determined by evaluating the sign of the eigenvalues obtained from matrix (5.8) and (5.9) as follows:

- If the eigenvalues of a matrix has all real parts less than zero, then the steady state is stable.
- If at least one of the eigenvalues has a real part greater than zero, then the

steady state is unstable.

- If at least one of the eigenvalues has a real part equals to zero, then no conclusion can be made about the stability.

To obtain the eigenvalues, κ , we solve the characteristic equation of matrix (5.8) and (5.9). The characteristic equation is obtained by calculating the determinant of $J_i - \kappa I$ (where I is the 3x3 identity matrix and i is either d_f or d) and equating it to zero i.e. $\det | J_i - \kappa I | = 0$. Therefore, the characteristic equation becomes a polynomial of order 3 such that:

$$\kappa^3 A + \kappa^2 B + \kappa C + D = 0 \quad (5.10)$$

Solving for κ generated from the matrix (5.8) gives the following eigenvalues:

$$\begin{aligned} \kappa_1 &= \frac{\alpha_0 b \omega}{\mu + \omega} - \mu - \gamma_0 \\ \kappa_2 &= -\mu \\ \kappa_3 &= -\mu \\ \kappa_4 &= -\mu - \omega \end{aligned} \quad (5.11)$$

From the set of Equ.(5.11) κ_2 , κ_3 and κ_4 have got all real parts less than zero since all of parameters μ and ω are always greater than zero. To check the sign of κ_1 , let us suppose that $Z = \frac{\alpha_0 b \omega}{\mu + \omega}$. We can then express R_0 as a function of Z . Thus $R_0 = \frac{\alpha_0 b \omega}{(\mu + \omega) \gamma_0 + \mu} = Z \frac{1}{\mu + \gamma_0}$. At the invasion threshold, $R_0 \geq 1$ and therefore $\frac{Z}{\mu + \gamma_0} \geq 1$. Making Z the subject gives us $Z \geq \mu + \gamma_0$ whenever $R_0 \geq 1$. Hence, when $R_0 \geq 1$, then $\kappa_1 > 0$ making the disease free equilibrium unstable and whenever $R_0 < 1$, then $Z < \mu + \gamma_0$ making $\kappa_1 < 0$ resulting to the disease free

equilibrium being stable to small perturbations.

Solving for κ generated from matrix (5.9) gives the following eigen values:

$$\begin{aligned}\kappa_1 &= -\mu \\ \kappa_2 &= 0 \\ \kappa_3 &= 0\end{aligned}\tag{5.12}$$

and

$$\begin{aligned}\kappa_{4(i)} &= ((12\gamma_0\mu^5 + 8\mu^5\omega + 4\mu^6 + 12\gamma_0^2\mu^4 + 4\gamma_0^3\mu^3 + 4\mu^4\omega^2 + 24\gamma_0\mu^4\omega \\ &\quad + 12\gamma_0\mu^3\omega^2 + 24\gamma_0^2\mu^3\omega + 4\gamma_0^3\mu\omega^2 + 8\gamma_0^3\mu^2\omega + 12\gamma_0^2\mu^2\omega^2 - 4\alpha_0b\mu^4\omega \\ &\quad + \alpha_0^2b^2\mu^2\omega^2 - 4\alpha_0b\mu^3\omega^2 - 8\alpha_0b\gamma_0\mu^2\omega^2 - 4\alpha_0b\gamma_0^2\mu\omega^2 - 4\alpha_0b\gamma_0^2\mu^2\omega \\ &\quad - 8\alpha_0b\gamma_0\mu^3\omega)^{\frac{1}{2}} - \alpha_0b\mu\omega)/(2(\gamma_0\mu + \gamma_0\omega + \mu\omega + \mu^2))\end{aligned}\tag{5.13}$$

if $\gamma_0\mu + \gamma_0\omega + \mu\omega + \mu^2 \neq 0$ and

$$\begin{aligned}\kappa_{4(ii)} &= (2\gamma_0\mu^2 + \gamma_0^2\mu + \gamma_0^2\omega + \mu^2\omega + \mu^3 + 2\gamma_0\mu\omega \\ &\quad - \alpha_0b\gamma_0\omega - \alpha_0b\mu\omega)/(\alpha_0b\omega)\end{aligned}\tag{5.14}$$

if $\gamma_0\mu + \gamma_0\omega + \mu\omega + \mu^2 = 0$ and $\alpha_0, b, \omega \neq 0$.

Since one of the eigenvalues from Eqn.(5.12) is zero, then no conclusion about the stability of the model can be made. Additionally, the equation set (5.13) and (5.14) quickly become analytically intractable on trying to determine the behaviour of κ . For this reason, the behaviour of the model beyond the invasion threshold is presented in the next section using numerical simulations.

5.3.3 Numerical simulations

The following section investigates the temporal behaviour of the homogeneous model and confirms some of the analytical results in section 5.3.1 using the numerical solution obtained from Eqn.(5.1). The integration of the system of ODE was done using MATLAB® [128] using the Runge-Kutta method of order 4,5 with an adaptive time step. During the simulations, the population was held constant by equating fertility and mortality rates. The parameters were held constant as shown in Table 5.1 unless otherwise stated and the initial conditions were set to $M=240,000$, $I_0 = 1$ and all the other state variables were set to 0.

Figure 5.2 shows the dynamic behaviour of the epidemiological compartments in the model presented in Figure 5.1. Initially, the model is started without infection and this leads to the disease free equilibrium where almost all the individuals remain in the S_0 class. Introduction of a single primary infectious person is done at time 20 years and the model attains the endemic equilibrium with damped oscillations. During the course of the epidemic, after the introduction of the infection in the population, the primary susceptibles gradually become infected and move into the I_0 class. Individuals then gradually move through the recovered P_0 class and into the second level of infection. Majority of the people in the population, at equilibrium, are in the S_2 class. This is expected since most of the people will have experienced their first and second infections and the duration of infection is much shorter in the I_1 and I_2 classes. At endemic equilibrium, the number of people in each state remains unchanged and the effective reproduction number denoted by R_e is unity [116, 10]. The effective reproduction number is defined as the number of new infections that a primary infected individual infects during their

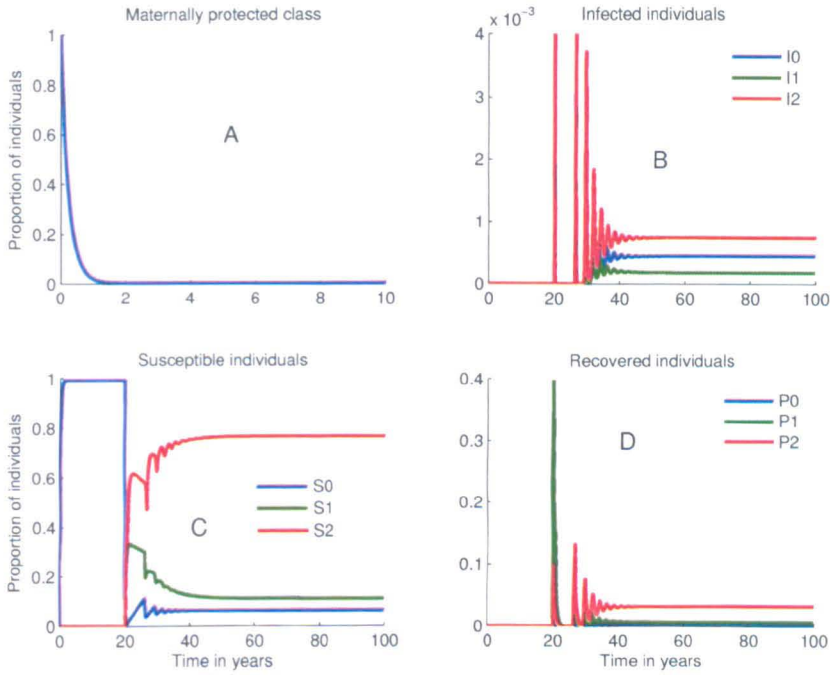


Figure 5.2: Shows the behaviour of the model compartments at both the disease free equilibrium and following the introduction of a single primary infected at time 20 years using the baseline parameters in Table 5.1. The y-axis shows the proportion of individuals in any of the compartments and the x-axis shows time in years. Subplot(A) shows the behaviour of the Maternally protected class, subplot(B) that of the infected classes, subplot(C) that of the susceptible and partially susceptible classes and subplot(D) that of the recovered classes.

infectious period. The model proceeds to dynamic equilibrium in an oscillatory manner which is typical of this family of models [116]. The value of the transmission parameter, b , used during the simulations was 300 giving a basic reproduction number of 7.36. This is in close range of the best fitting MSEIRS4 (Maternally protected Susceptible Exposed Infected Recovered Susceptible) model explored by Weber et al [201] and White et al [205] whose R_0 values range between 5.3-7.2 and 9.2-9.4 respectively.

We will use numerical simulations to validate the results obtained in Eqn.(5.4) for the invasion threshold. Given that we have the baseline parameters in Table 5.1, we will replace the values of μ , γ_0 , ω and α_0 into Eqn.(5.4). Replacing these values gives us an invasion threshold, $b^* = 40.778$. To compare this value with the numerical estimate, we will plot the proportion of infected individuals in the I_0 , I_1 and I_2 classes in the population at equilibrium as a function of the transmission parameter. Figure 5.3 shows the resulting invasion threshold. At the invasion threshold the model indicates a qualitative behaviour at which point the system branches from the disease free equilibrium to the endemic equilibrium. The point at which this happens is known as the bifurcation point and the resulting curve is referred to as the bifurcation curve. The parameter on the x-axis, b , is referred to as the bifurcation parameter.

From Figure 5.3, the bifurcation point is estimated to be 40.6 which is in agreement with the analytic solution. The reason the two values are not equal is because, in the numerical solution the step size of b was taken to be 0.2 and therefore the next value from 40.6 was 40.8 which is above the threshold. The numerical solution can be improved by reducing the incremental step of the transmission parameter to the desired accuracy. The refining of the incremental step would however demand a reasonable amount of computer time and therefore it was not done. The kind of bifurcation curve shown in Figure 5.3 is known as a forward bifurcation curve since the nature of the curve is such that as we travel along it beyond the bifurcation point, the level of infection increases as the bifurcation parameter increases [74].

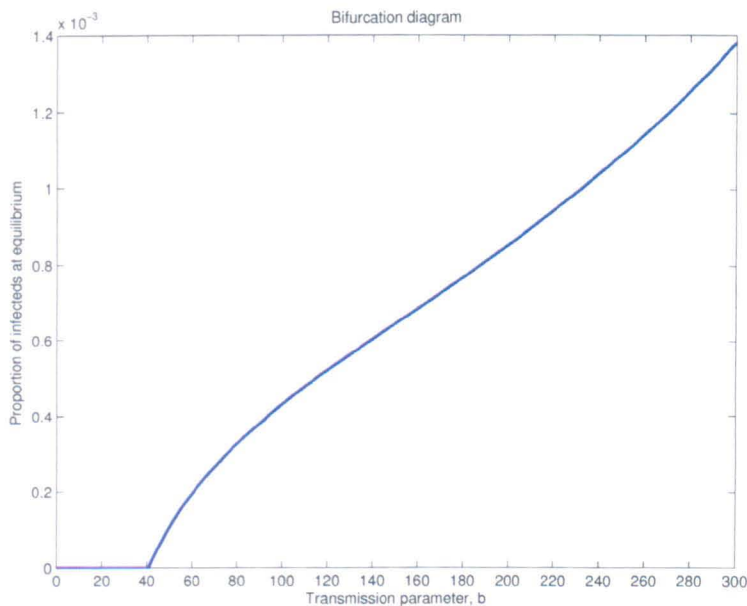


Figure 5.3: Bifurcation diagram showing the invasion threshold. The transmission parameter on the x-axis is the bifurcation parameter while the y-axis shows the proportion of infecteds i.e. I_0 , I_1 and I_2 in the population at equilibrium

5.3.4 Sensitivity analysis

By changing the model parameters in a controlled manner and observing the effect they have on the output gives us an idea of how robust the model is to changes in the input. We will carry out the sensitivity analysis by investigating how the model input parameter values will affect the model's output and especially the kind of bifurcation curve it exhibits.

In order to carry out the analysis, we identified parameters that we had little information about and varied them within bounded ranges. We identified the following parameters to constitute the sensitivity analysis: $\sigma_0, \sigma_1, \sigma_2$ which represent the partial immunity factor, η_0, η_1, η_2 representing the partial immunity factor

reducing susceptibility and $\alpha_0, \alpha_1, \alpha_2$ as factor reducing infectiousness of infected individuals. The simulations in Figure 5.4 were run with the initial conditions similar to that shown in Figure 5.2 but with small changes in σ_i, η_i and α_i where $i = 0, 1, 2$. From figure 5.4, we can see that the equilibrium proportion infected in the population increases with an increase in infectivity of individuals and with increased susceptibility. One can see this sustained increase in the equilibrium proportion infected from subplot 1 to 6 and subplot 7 to 12. The bifurcation points for all the parameter combinations have remained at 40.6.

Alternatively, instead of starting with the initial conditions where everyone is in the maternally protected class and a single person in the infected, I_0 , class, we started with a more general situation where the number of individuals in the epidemiological classes are distributed in a more realistic way but not necessarily at equilibrium such that $S_0 < S_1 < S_2$, $I_0 < I_1, I_2$ and $P_0 < P_1 < P_2$. The bifurcation diagrams, not shown, obtained using the new initial conditions were exactly the same as those obtained in Figure 5.4. So far, the algorithm that we have used to produce the bifurcation plots may only be useful in producing curves exhibiting forward bifurcation. This is so thought because the procedure did not succeed in producing a backward bifurcation for a model with a well known set of parameters resulting in multiple sub-critical endemic equilibria [79].

As a result, we adopted the following heuristic procedure: we started by evolving the system through time and plotting the equilibrium proportion of infected people in the population for every value of the bifurcation parameter. In this case the parameter is incremented from 0 to 500. We then proceeded by simulating the model through time a second time but now decrementing the bifurcation parameter from 500 to 0. The initial conditions for each "backward" step was taken to be

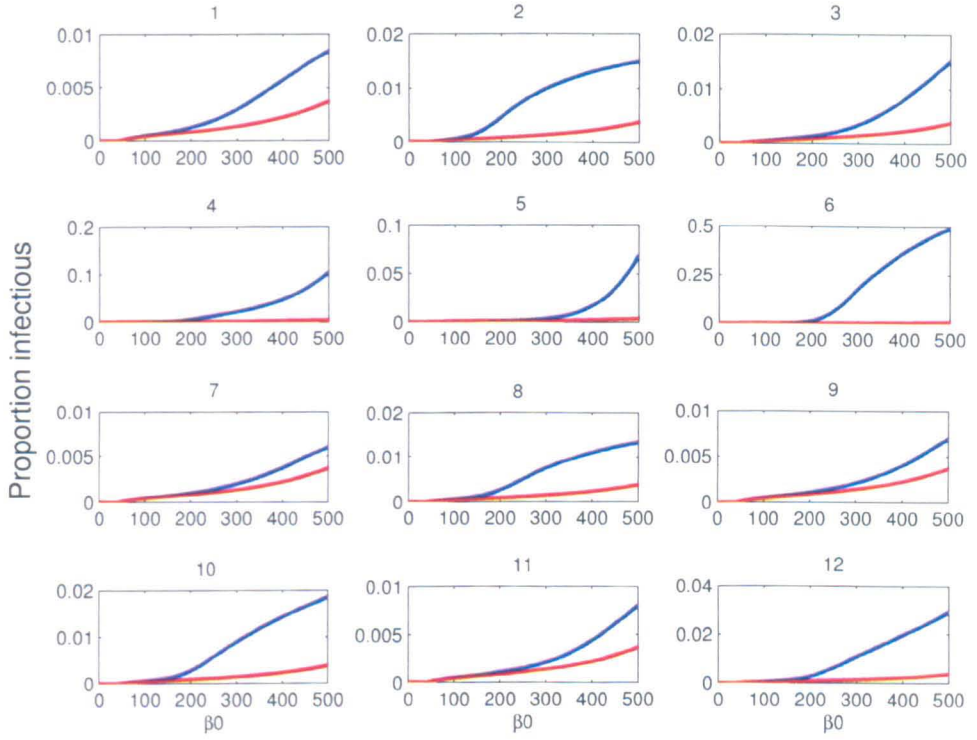


Figure 5.4: Sensitivity analysis depicting the different parameter values used. The following parameter estimates were used for the graphs above. Subplot 1 and 7 $\eta_0 = \eta_1 = \eta_2 = 0$ and $\alpha_0 = 1, \alpha_1 = \alpha_2 = 0.5$ subplot 2 and 8 $\eta_0 = \eta_1 = \eta_2 = 0$ and $\alpha_0 = \alpha_1 = \alpha_2 = 1$ subplot 3 and 9 $\eta_0 = \eta_1 = \eta_2 = 0.5$ and $\alpha_0 = 1, \alpha_1 = \alpha_2 = 0.5$ subplot 4 and 10 $\eta_0 = \eta_1 = \eta_2 = 0.5$ and $\alpha_0 = \alpha_1 = \alpha_2 = 1$ plot 5 and 11 $\eta_0 = \eta_1 = \eta_2 = 1$ and $\alpha_0 = 1, \alpha_1 = \alpha_2 = 0.5$ subplot 6 and 12 $\eta_0 = \eta_1 = \eta_2 = 1$ and $\alpha_0 = \alpha_1 = \alpha_2 = 1$. For subplot 1 to 6 $\sigma_0 = 0.7, \sigma_1 = \sigma_0^2, \sigma_2 = \sigma_0^3$ and subplot 7 to 12 $\sigma_0 = 0.7, \sigma_1 = \sigma_0 - 0.2, \sigma_2 = \sigma_0 - 0.4$. The red curve shows the base case as shown in Figure 5.3.

the final equilibrium values of the system at the point where the bifurcation parameter is at its maximum. We tested this procedure against a set of models with well known parameter sets resulting to backward bifurcation and the procedure produced the desired results. However, this does not mean that it will work with

all kind of models but it gives us confidence to apply the procedure to our model. We applied the procedure to the RSV model but it remained quite robust to the invasion threshold in Figure 5.3. All the parameter values chosen resulted to a model displaying forward bifurcation. Both the curves resulting from the forward run and the backward run perfectly overlapped.

We then introduced vaccination in the model to explore what effect it would have on the bifurcation curve observed. This was inspired by previous work done by Greenhalgh et al [74, 75, 73] on a two stage and three stage bovine RSV models. Vaccination was assumed to move newborn individuals from the M class to the S_2 class. This does not necessarily represent reality since it is unlikely that the vaccine will provide better immune response than a primary infection. Using this model, our previous calculation of the invasion threshold and the basic reproduction number does not hold. We therefore re-calculate the basic reproduction number by applying the procedure explained in [195, 42]. The resulting R_0 is given by Eqn.(5.15).

$$R_0 = \frac{\alpha_0 b \omega}{(\mu + \omega)(\gamma_0 + \mu)} + \left(1 + \frac{\gamma_0}{(\mu + \omega)(\mu + \gamma_0)(\gamma_2 + \mu)}\right) b \mu \sigma_1 \alpha_2 \phi \quad (5.15)$$

where ϕ represents the proportion of individuals effectively vaccinated and protected since it is a compound parameter representing the vaccine efficacy and vaccine uptake. We have also assumed that the force of infection acting on the S_1 and S_2 classes is different from that acting on the primary susceptibles, S_0 . We therefore introduced three new transmission parameters β_0, β_1 and β_2 which are indicative of the difference in the transmission potential as individuals gain experience by being infected multiple times. Infectivity may not be directly or

easily measured and therefore in this modified model we have assumed that it is a compound parameter within the transmission parameter and that allows us to set the parameters α_0 , α_1 and α_2 to 1. The new force of infection function can therefore be expressed as:

$$\lambda = \frac{1}{N} \sum_{i=0}^2 \beta_i I_i \quad (5.16)$$

where $\beta_i = b_i \alpha_i$.

To plot the bifurcation diagram, we used the parameter β_0 as the bifurcation parameter holding all the other parameters constant during the simulations. Figure 5.5 shows the resulting bifurcation diagram with parameter values indicated in Table 5.2 in the column labelled baseline. The resulting bifurcation diagram is known as a backward bifurcation curve. This is because there exists multiple super-critical endemic equilibria. From the Figure, the discontinuity shows that there exists two endemic equilibrium in the interval $45 \leq \beta_0 \leq 55$. The sensitivity analysis column in Table 5.2 shows the range of the parameter values where the backward bifurcation is conserved. The regions were determined by holding all other parameters constant to values in the baseline column and then varying a single parameter in the model.

Table 5.2: Parameter estimates used to produce the backward bifurcation diagram in Figure 5.5 (column labelled baseline) and the parameter space within which the backward bifurcation is conserved (column labelled sensitivity analysis)

Parameter symbol	Description	Baseline	Sensitivity analysis
μ	Death/birth rate	0.02/yr	$0.005 \leq \mu \leq 0.2$
ω	Rate of decay of maternal antibodies	4/yr	$1 \leq \omega \leq 53.143$
β_0	Transmission parameter (Primary infecteds)	Used as the bifurcation parameter	
σ_i	Long-term immunity factor reducing the susceptibility of previously exposed individuals in S_i and P_i	$\sigma_0 = 1$ $\sigma_1, \sigma_2 = 0.5$	$0.611 \leq \sigma_0 \leq 1$ $0.497 \leq \sigma_1 \leq 1$ $0 \leq \sigma_2 \leq 1$
ρ_i	Rate of waning of short-term immunity of recovered individuals, P_i $i=0,1,2$	$\rho_0 = 4$ $\rho_1, \rho_2 = 2$	$0.01 \leq \rho_0 \leq 52.143$ $0.28 \leq \rho_1 \leq 52.142$ $1.157 \leq \rho_2 \leq 52.143$
η_i	Short-term immunity factor additionally reducing the susceptibility of recovered individuals, P_i	$\eta_0 = 1$ $\eta_1, \eta_2 = 0.5$	$0 \leq \eta_0, \eta_1, \eta_2 \leq 1$
γ_i	Rate of recovery from infected classes I_i into the P_i classes	$\gamma_0 = 36.5$ $\gamma_1 = 40$ $\gamma_2 = 40$	$16 \leq \gamma_0 \leq 138.732$ $16 \leq \gamma_1 \leq 180$ $16 \leq \gamma_2 \leq 39.88$
α_i	Factor reducing infectiousness of I_0, I_1 and I_2	$\alpha_0, \alpha_1, \alpha_2 = 1$	$0.267 \leq \alpha_0 \leq 1$ $0 \leq \alpha_1 \leq 1$ $0.98 \leq \alpha_2 \leq 0$
ϕ	Effective vaccination coverage	0	$0 \leq \phi \leq 1$
β_1	Transmission parameter (Secondary infecteds)	10.95	$0 \leq b_1 \leq 1000$
β_2	Transmission parameter (Tertiary infecteds)	86.87	$86.364 \leq b_2 \leq 1000$

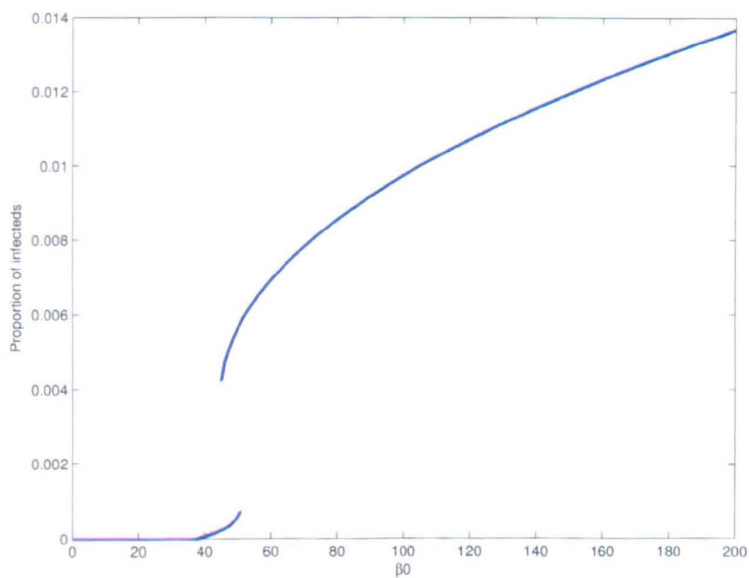


Figure 5.5: Shows the proportion infectious at equilibrium. The x-axis shows β_0 which is used as the bifurcation parameter and the rest of the parameters are as shown in Table 5.2 in the column labelled baseline.

5.4 Summary and discussion

In this chapter, I have presented and described a simple epidemiological model describing the transmission of Respiratory Syncytial Virus (RSV). I have used a deterministic model in section 5.2 with 10 different epidemiological compartments. The comparative simplicity of the model allows for a partial stability analysis to be carried out as well as the calculation of the analytical invasion threshold. However, even with such a simple model, investigating the stability of the endemic equilibrium is mathematically difficult. Investigating the equilibrium properties reveal that the disease free equilibrium is locally stable provided that $R_0 < 1$ and locally unstable provided that $R_0 \geq 1$. The model was solved numerically to test and confirm the analytical results and at the same time to explore the dynamic behaviour at equilibrium. The model achieves dynamic equilibrium with damped oscillations as can be seen in Figure 5.2.

The model without vaccination consistently generated forward bifurcation diagrams for different parameter values and initial conditions. However, in the presence of vaccination, the model displays backward bifurcation with multiple supercritical endemic states. Figure 5.5 was produced with a parameter set in Table 5.2 in the column labelled baseline. This implies that it is possible to have this kind of backward bifurcation for what may be assumed to be realistic parameter values for RSV transmission. However, it must be admitted that there might be some issues in the epidemiological interpretation of some of the parameters. For example, the inclusion of the extra transmission parameters β_0, β_1 and β_2 makes the parameters collinear with the α_i , $i = 0, 1, 2$, and therefore making the α_i^s redundant. The reason is because α_i^s modify the infectiousness of the infected classes and with the

introduction of β_i^s the infectiousness can now be modified by β_i^s alone since β_i^s become a compound parameter inherently accounting for infectiousness. Within the sensitivity analysis region, β_2 can potentially be greater than β_1 . This may be construed to mean that susceptible individuals who have had three or more infections are more susceptible to infection than those who have had a single previous infection. This seems contrary to previous studies reporting a decline in the attack rate by experience of infection [93, 68]. The sensitivity analysis reported in this chapter gives a parameter space region within which the backward bifurcation is conserved. The univariate sensitivity analysis has the advantage that it is quick and simple since you hold $k - 1$ parameters constant and only vary one parameter at a time. However, the disadvantage of this approach is that only a small region of the k -dimensional parameter space is evaluated and the remaining parameters have to be estimated with a high degree of certainty. There are however more sophisticated statistical techniques that have been developed to allow the simultaneous variation of different parameters. One such technique, Latin Hypercube Sampling, has been considered in Chapter 6.

The presence of backward bifurcation has got certain implications for infection control. Classical epidemiological modelling suggests that at the invasion threshold R_0 should be equal to 1. However, this is not always the case for a model displaying backward bifurcation with multiple sub-critical endemic equilibrium points as the condition is no longer satisfied. In order for the infection to be eliminated from the study population, we would require that $R_0 < \dot{R}_0$ where \dot{R}_0 is the value of R_0 corresponding to a vertical turning point on the bifurcation curve. So, a model displaying backward bifurcation with multiple sub-critical endemic equilibria would require more effort to eliminate the infection in the population compared

to a model displaying forward bifurcation.

However, the discussion of the results should be looked at in the light of the limitations of the model presented. The analysis presented is based on a deterministic model and hence the results will be valid under a large population size. If the spread of the infection is within a small population, then a stochastic model may be more appropriate.

Chapter 6

Modelling RSV transmission dynamics and the potential impact of vaccination

6.1 Introduction

Mathematical models have previously been used to inform decisions on the impact of vaccination on the transmission of infectious diseases [147, 125, 47, 164]. Vaccination has not only been shown to be an extremely effective way of controlling infections in vaccinated individuals but also in those not vaccinated through indirect protection [124, 66]. This mechanism of indirect protection has been attributed, in part, to the eradication of some infectious diseases e.g. small pox [58]. A proper evaluation of direct and indirect effects of vaccination need to be evaluated before the introduction of a vaccine in the population. RSV has no licenced vaccine although there are some under development that are promising (see the review in Chapter 2), and this would be the opportune time to evaluate the optimal allocation of such a vaccine and to evaluate the optimal age at which

to vaccinate taking into consideration the indirect effect of vaccination. A series of mathematical models to describe the transmission of RSV and the effect of introducing an RSV vaccine would therefore be a valuable tool for exploring for example the optimal vaccination strategy.

In this chapter, we have presented an investigation of the epidemiology of RSV and the potential outcome of routine vaccination using a realistic age structured model. We have generalized the mathematical model developed in Chapter 5 to include age structure and the age specific risks of developing disease given an infection and risk of hospitalization given disease. The model parameters are derived from published studies. The model has been validated against RSV hospitalization data from Kilifi District Hospital (KDH). Transmission rates between individuals are determined by the Who Acquires Infection From Whom (WAIFW) matrix [151, 10] which has been parameterized using self reported rates of social contact from a diary study [196, 136] and a synthetic contact mixing matrix, see Chapter 4. The characterization of the social contact data in computational mathematical models is important since it has been shown that the transmission potential of an infection is strongly dependent on mixing patterns between individuals which in turn depends on the socio-demographic parameters [173].

Using the fully parameterized RSV transmission model, we then implement routine vaccination assuming two types of vaccines: one that acts to prevent primary infections and the other that works to prevent all infections. Given that the results of any modelling process are confounded by the presence of uncertainties, we have presented a global uncertainty and sensitivity analysis to assess how robust the model outcomes are to different assumptions. The sensitivity analysis has been carried out using Latin Hypercube Sampling (LHS) [129] and the method of

Partial Rank Correlation Coefficient (PRCC).

6.1.1 Objectives

The objectives of the work presented in this chapter are:

- to develop a Realistic Age-Structured (RAS) mathematical model that accurately reflects the transmission of RSV within a community of individuals
- to evaluate the impact of routine vaccination on the number of RSV hospitalizations and identify the optimal age at which to vaccinate given the two mixing assumptions presented in Chapter 4
- to assess the variability in the outcome of interest that is due to uncertainty in estimating the values in the input parameters using LHS sampling and PRCC.

6.2 Methods

6.2.1 Model structure

The mathematical model developed aimed at simulating the transmission dynamics of RSV in an age-structured population is presented. The model is subdivided into two sub-models, one capturing the demographics while the other one the epidemiological dynamics of RSV.

6.2.1.1 Demographic sub-model

The model draws from previously published age-structured models [96, 180, 164] making use of discrete age classes rather than continuous age classes requiring the use of Partial Differential Equations (PDE). A system of PDE is difficult to handle numerically since it requires time consuming computer simulations while on the other hand systems of ODEs are relatively easy to solve, parameterize and easier to conceptualize. The demographic sub-model is divided into 99 age classes i.e. 24 monthly age classes in the first two years of life and yearly age classes from the third year of life. Individuals older than 76 years have been put together in the final age class. The selection of monthly age classes is chosen so as to capture the transmission dynamics and the impact of vaccination in the most critical age groups. We relax the assumption of a constant population by making use of the age-specific mortality and fertility rates from the KHDSS since the demographic sub-model is intended to correspond to its population structure and growth. Temporal changes in the mortality and fertility rates have not been included since their inclusion would possibly make the demographic model more complicated and difficult to tell whether features observed in the model are as a result of demographic, epidemiological or vaccination patterns. The number of people in each age class is allowed to vary as a result of a continuous ageing process from a younger to an older age group and through natural deaths. The rate of ageing is taken to be the reciprocal of the length of the source age class. Assuming that an age class can be represented as $[a_i, a_{i+1}]$, then the rate of ageing from age class i to $i + 1$, denoted as κ_i , will be expressed as $\kappa_i = \frac{1}{a_{i+1} - a_i}$. Describing the ageing process using a rate results in a fixed proportion of individuals in an

age class moving into the next class at each time step with exponential duration of stay and with duration equal to the length of the age class. However, having several age classes within the first two years of life minimizes the limitation of having exponentially distributed time since the duration of stay now approaches a gamma distributed time for the first and the second years of life.

6.2.1.2 Epidemiological sub-model

As in Chapter 5, the host population is stratified into 10 epidemiological groups: those that have maternal protection (M), primary susceptibles (S_0), primary infecteds (I_0), primary recovered (P_0), secondary susceptibles (S_1), secondary infecteds (I_1), secondary recovered (P_1), tertiary susceptibles (S_2), tertiary infecteds (I_2) and tertiary recovered (P_2). We have modelled several stages within the M class so as to include a more realistic distribution of waning of RSV specific maternal antibodies. The diagram in Figure 6.1 shows the flow of individuals through the epidemiological compartments. A notable difference between the model structure presented in Chapter 5 and this one is that immunity to RSV is now considered to be temporary but solid while in the previous one we had a combination of both temporary and partial immunity. The partial immunity was dropped due lack of sufficient support information from published data and was only included on an exploratory basis in Chapter 5. However, there exists evidence from recently published data that individuals remain protected while in the P class showing that over a six month period, following infection, an individual has about 70% protection from infection [155]. Data from literature also suggests that individuals rarely get infected more than once during the same epidemic [3] and hence we have modelled solid immunity for approximately six months. Sensitivity and uncertainty

analysis presented later explores what effect this assumption has on the number of hospitalizations and the optimal age at vaccination. The number of M classes is determined by fitting the model to time series and age-specific RSV related hospitalizations in the KDH (Kilifi District Hospital) which will be presented later. As the host population is now stratified by age, the model includes age-dependent processes, such as the force of infection. Thus all the state variables are stratified by both age and time such that $S_{0,i}(t)$ represents the density of the primary susceptibles of age i at time t and so forth. The rates, with respect to both time and age, at which individuals flow from one epidemiological state to another are described in the system of ordinary differential equation shown Eqn.(6.1): where q is the number of M sub-classes, κ_i is the rate of ageing, Θ_i is the fertility rate, $\gamma_{0,i}$, $\gamma_{1,i}$ and $\gamma_{2,i}$ are the recovery rates from primary, secondary and tertiary infections respectively, ω_i is the rate of loss of maternal antibodies, $\rho_{1,i}$ and $\rho_{2,i}$ are the rates of loss of secondary and tertiary immunity respectively, λ_i is the per capita rate of infection, N_i is the total number of people in age class i and μ_i is the mortality rate where i represents the age class. We have ignored RSV related deaths since, relative to all other causes of death (even in childhood), RSV related mortality is negligible and therefore they are not represented in Eqn.(6.1). For a list of the model parameters and their source, see Tables 6.1, 6.2 and 6.3.

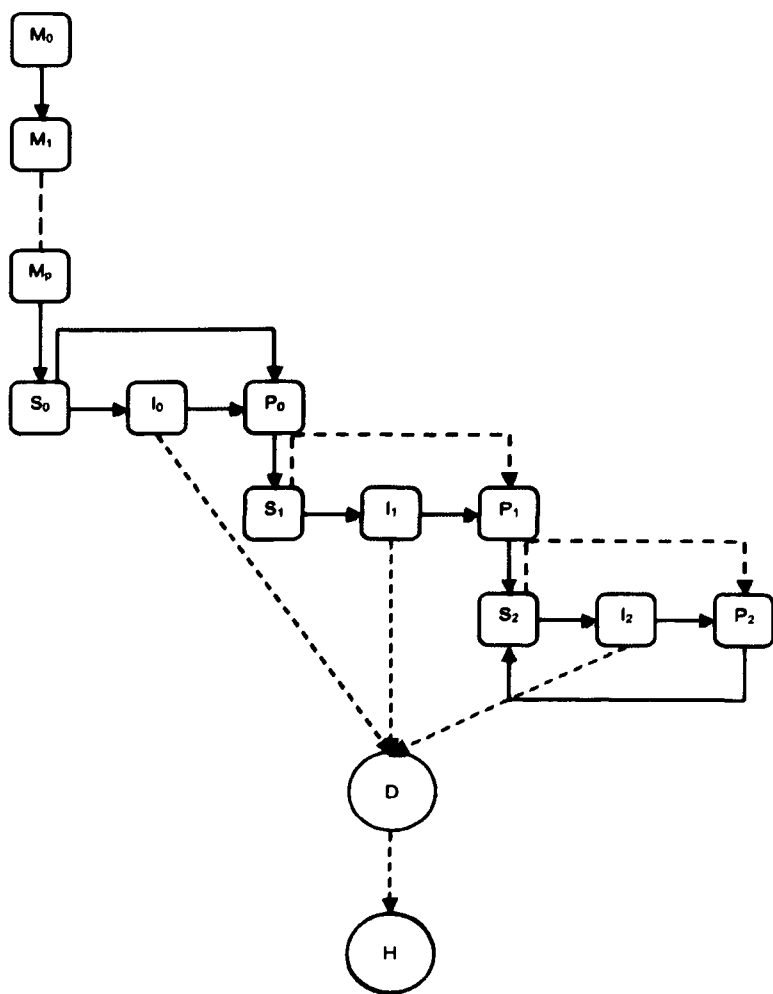


Figure 6.1: Schematic flow diagram of the compartmental model. Arrows represent the flow of individuals between states.

If $q = 1$

$$\frac{dM_{0,i}}{dt} = \kappa_{i-1}M_{0,i-1} + \begin{bmatrix} 0 & \text{if } i \neq 1 \\ \sum_{j=1}^n \Theta_j N_j & \text{if } i = 1 \end{bmatrix} - M_{0,i}(\mu_i + \omega_i q) - \kappa_i M_{0,i}$$

else $q \geq 2$

$$\begin{aligned} \frac{dM_{q-1,i}}{dt} &= \kappa_{i-1}M_{q-1,i-1} + \omega_i q M_{q-1,i} - \kappa_i M_{q-1,i} - \mu_i M_{q-1,i} \\ \frac{dS_{0,i}}{dt} &= \kappa_{i-1}S_{0,i-1} + \omega_i q M_{q-1,i} - S_{0,i}(\mu_i + \lambda_i) - \kappa_i S_{0,i} \\ \frac{dI_{0,i}}{dt} &= \kappa_{i-1}I_{0,i-1} + \lambda_i S_{0,i} - I_{0,i}(\mu_i + \gamma_{0,i}) + \kappa_i I_{0,i} \\ \frac{dP_{0,i}}{dt} &= \kappa_{i-1}P_{0,i-1} + \gamma_{0,i}I_{0,i} - P_{0,i}(\mu_i + \rho_{0,i}) - \kappa_i P_{0,i} \\ \frac{dS_{1,i}}{dt} &= \kappa_{i-1}S_{1,i-1} + \rho_{0,i}P_{0,i} - S_{1,i}(\mu_i + \sigma_{0,i}\lambda_i) - \kappa_i S_{1,i} \\ \frac{dI_{1,i}}{dt} &= \kappa_{i-1}I_{1,i-1} + \sigma_{0,i}\lambda_i S_{1,i} - I_{1,i}(\mu_i + \gamma_{1,i}) - \kappa_i I_{1,i} \\ \frac{dP_{1,i}}{dt} &= \kappa_{i-1}P_{1,i-1} + \gamma_{1,i}I_{1,i} - P_{1,i}(\mu_i + \rho_{1,i}) - \kappa_i P_{1,i} \\ \frac{dS_{2,i}}{dt} &= \kappa_{i-1}S_{2,i-1} + \rho_{1,i}P_{1,i} + \rho_{2,i}P_{2,i} - S_{2,i}(\mu_i + \sigma_{1,i}\lambda_i) - \kappa_i S_{2,i} \\ \frac{dI_{2,i}}{dt} &= \kappa_{i-1}I_{2,i-1} + \sigma_{1,i}\lambda_i S_{2,i} - I_{2,i}(\mu_i + \gamma_{2,i}) - \kappa_i I_{2,i} \\ \frac{dP_{2,i}}{dt} &= \kappa_{i-1}P_{2,i-1} + \gamma_{2,i}I_{2,i} - P_{2,i}(\mu_i + \rho_{2,i}) - \kappa_i P_{2,i} \end{aligned} \tag{6.1}$$

6.2.1.3 Force of infection

The force of infection is defined as the per person rate at which a susceptible individual becomes infected. The force of infection in the model is both time and

age-specific and is denoted by λ_i . Determining the age-specific force of infection is problematic and difficulties arise because there are usually more unknown age-specific transmission parameters than observations on risk of infection for each age class and as a consequence we face an indeterminacy problem [180, 54]. Often, a-priori assumptions about contact processes are invoked in order to reduce the number of unknowns to the number of age classes [56, 10]. The simplest such hypothesis is homogeneous mixing which is assumed in the model developed in Chapter 5. To overcome the problem of assuming that the mixing pattern is known a-priori, we can infer a likely contact pattern from detailed household occupancy data and from social contact data reported from paper diaries [136, 48]. Two contact matrices have been developed as part of this work, refer to Chapter 4, and have been applied to the current age specific RSV model. Therefore to calculate the force of infection we make use of the *social contact hypothesis* suggested by Wallinga et al [196] which states that the WAIFW matrix (age-specific transmission parameters) is proportional to the number of self-reported age-specific social contacts such that $\beta_{i,j} = q_c C_{i,j}$ where $C_{i,j}$ is the social contact matrix, $\beta_{i,j}$ is the WAIFW matrix, q_c is a proportionality factor that measures the disease specific infectivity and i, j are the age classes. Seasonality of RSV is included using a cosinusoidal function since the main drivers are not well understood. Seasonality of RSV is a more complicated process that is probably affected by a combination of a changing contact pattern, changing immunity and a constantly changing virus. The age-specific force of infection is therefore expressed as shown in Eqn.(6.2).

$$\lambda_i(t) = \sum_{j=1}^n \left(\frac{\beta_{i,j} (1 + a \cos(2\pi(t - \phi)))}{N_i(t)} \sum_{k=0}^2 \alpha_k I_{k,j}(t) \right) \quad (6.2)$$

where $\beta_{i,j}$ represents the transmission coefficient between susceptibles of age i and infecteds of age j , $N_i(t)$ is the total number of individuals in age class i at time t , α_0, α_1 and α_2 represent the relative infectiousness of infected individuals in I_0, I_1 and I_2 classes respectively. The seasonal parameters defining the relative amplitude, a , and the peak in transmission, ϕ , are unknown and are determined by fitting the model to RSV specific hospitalization data from the Kilifi District Hospital.

The initial conditions i.e. the values of all the state variables for each age class at time $t = 0$ and the boundary conditions i.e. the values of the state variables for the first age class are taken to be the pre-vaccination numbers found by running the model to its stable limit cycle.

6.2.2 Parameter estimates and model fitting

Many of the model parameters that have been identified from literature can be found in Chapter 3 while the contact data that has been used to parameterize the force of infection can be found in Chapter 4. Tables 6.1, 6.2 and 6.3 give a summary of the parameter estimates that have been used in the numerical simulations i.e. their baseline value and the source. A parameter that is worth highlighting at this point is the factor reducing the infectiousness of I_1 and I_2 denoted by α_1 and α_2 respectively. Hall et al [85] suggested that infants shed the virus in large quantities and for prolonged periods of time. The large quantities shed can equally be attributed to primary infections since these infants are most likely experiencing their first infection. Reduction in shedding quantities is reported to be a function of age with reported decline in shedding quantities with increasing age.

This means that the quantity of virus shed will also decrease with the number of previous RSV infections. We have therefore assumed that second infections are half as infectious as primary infections and third and subsequent infections are half as infectious as second infections. We have therefore estimated α_1 and α_2 to be 0.5 and 0.25 respectively. These parameters have however been included in the uncertainty and sensitivity analysis to assess how different assumptions affect the output of interest.

The model predicts the number of individuals in the population in each state variable and at each time interval for each age class i where $i = 1, 2, \dots, 99$ while the hospitalization data gives the number of new hospitalized cases per month per age class. We therefore must use a scaling factor on the incidence of new infections in order to compare the model output with the hospitalization data. From Figure 6.1, the risk of developing disease following infection is shown by the circular compartment labelled D while the risk of hospitalization following development of disease is given by the compartment labelled H . The risk of development of disease given a primary, secondary and tertiary infection is given by the parameters d_0, d_1 and d_2 respectively (see Table 6.2) while the age-specific risk of hospitalization given that one has disease is given by parameter h (see Table 6.3). The age-specific risk of hospitalization was estimated by fitting the static model (model with force of infection that was age-specific but constant in time) to the age specific hospitalization data from the Kilifi District Hospital. The initial vector used was from published work [149] and there was reason to believe that the study underestimated the proportion of hospitalizations and hence the fitting to data. The output from the compartment labelled H in Figure 6.1 is what

we fit to the hospitalization data and can be expressed as shown in Eqn.(6.3).

$$H_i = h_i \lambda_i (d_{0,i} S_{0,i} + d_{1,i} \sigma_{1,i} S_{1,i} + d_{2,i} \sigma_{2,i} S_{2,i}) \quad (6.3)$$

where λ_i is the force of infection as shown in Eqn.(6.2). In fitting the model, we have used maximum likelihood estimation (MLE) method [207, 144] and given that hospitalizations is count data, we have assumed it follows a poisson distribution. MLE was chosen because it has several optimal properties: sufficiency (complete information about the parameter of interest is contained in its MLE estimator), consistency (true parameter value that generated the data is recovered asymptotically), efficiency (lowest possible variance achieved asymptotically) and parameterization invariance (same MLE estimate independent of the parameterization used) [144]. The negative log-likelihood was calculated as shown in Eqn (6.4) such that

$$LL = - \sum_{n=1}^{n_T} \sum_{i=1}^T \left(k_i \log \Lambda_i - \Lambda_i + \sum_{j=1}^{k_i} \log j \right) \quad (6.4)$$

where T is the number of observations for each month n , n_T is number of months and Λ_i is the corresponding expected incidences. To calculate the 95% confidence interval of the fitted parameters, we compute the central finite difference approximation to the Hessian of the log-likelihood estimates given the observed data to generate an asymptotic covariance matrix [121]. Using the covariance matrix, we compute the confidence interval using the normal approximation. However, the 95% confidence intervals calculated from the application of this procedure are expected to be narrow possibly due to the assumption of independent Poisson

observations for the likelihood calculations.

Table 6.1: Baseline parameter estimates used in the numerical simulations with the values based on the review in Chapter 3

Parameter symbol	Description	Baseline value	Data source
ω	Rate of decay of maternal antibodies		Fitted
$\beta_{i,j}$	Age-specific transmission parameter (WAIFW)	Fitted - Estimated by fitting contact matrices.	
σ_i	Long-term immunity factor reducing the susceptibility of previously exposed individuals in S_1 and S_2	$\sigma_1 = 0.75$ $\sigma_2 = 0.65$	[93]
ρ_i	Rate of waning of short-term immunity of recovered individuals, P_i $i=0,1,2$	$\rho_0 = 2yr^{-1}$ $\rho_1 = 2yr^{-1}$ $\rho_2 = 2yr^{-1}$	[3, 184]
γ_0	Rate of recovery from primary infection, I_0	$40.6yr^{-1}$	[87, 199]
γ_1, γ_2	Rate of recovery from secondary and tertiary infections, I_1, I_2	$93.7yr^{-1}$	[158, 87]
α_i	Factor reducing infectiousness of I_0, I_1 and I_2	$\alpha_1 = 0.5$ $\alpha_2 = 0.25$	See text for the justification
N	Initial population size	$\approx 240,000$	KHDSS
κ_i	Rate of ageing $0- \leq 24m$ $3-77yrs$	$12yr^{-1}$ $1yr^{-1}$	Reciprocal of the length of the age class
a	Relative amplitude		Fitted
ϕ	Phase angle - Peak of transmission		Fitted

Table 6.2: Baseline parameter estimates for the age-specific disease risk i.e. d_0, d_1 and d_2 following infections

Age in months	Primary infection	Second and subsequent infections	Data source
	$d_0(\%)$	$d_1, d_2(\%)$	
0-2	31.2	5.0	[155]
3-5	28.6	5.0	
6-8	20.0	5.0	
9-11	13.0	5.0	
12-17	7.6	5.0	
18-23	2.0	5.0	
≥ 24	2.0	2.0	

6.2.3 Vaccination

We assume that a vaccine will confer a protective effect which is equivalent to a natural infection. Therefore vaccination moves individuals from S_i to P_i where $i = 0, 1, 2$. We will consider two vaccine types: a vaccine that works to prevent primary infections (solid green line in Figure 6.1) and a vaccine that works against all infections (all green lines). Routine vaccination is implemented as individuals pass a defined age class. For example, implementing 80% vaccination coverage at at 2 months involves effectively vaccinating 80% of the susceptible individuals who are ageing from the first age class into the second one and moving them into their respective recovered classes i.e. $P_i : i = 0, 1, 2$. Vaccination provides both direct protection to those who are successfully immunized with the vaccine and indirect protection for those who are not immunized, by decreasing the likelihood that they will come into contact with an infectious individual. To calculate the effect

Table 6.3: Baseline parameter estimates for the age-specific risk of hospitalization following disease

Age in months	Risk of hospitalization, h (%)	Data source
1	32.76	
2	33.07	
3	21.9	
4	20.74	
5	18.86	
6	12.27	
7	9.4	
8	10.76	
9	9.1	
10	12.11	
11	9.87	Estimated by fitting the static model to age specific hospitalization data. The initial age-specific risk of hospitalization was sourced from [149]
12	6.7	
13	7.11	
14	7.78	
15	7.34	
16	4.13	
17	4.1	
18	10.52	
19	20	
20	14.84	
21	13.95	
22	10.18	
23	2.36	
24	8.41	
25-36	3.76	
37-48	1.08	
49-60	0.19	

of indirect protection, we run the model with a time invariant force of infection whose value is fixed to the pre-vaccination equilibrium.

6.2.4 Uncertainty and Sensitivity analysis

Due to the structural complexity coupled with a high degree of uncertainty in some of the model input parameters, the behaviour of the model to changes in these parameters is investigated using global uncertainty and sensitivity (U&S) analysis. Several approaches to this analysis exist ranging from a full factorial method to more sophisticated statistical methods allowing for simultaneous variation of the model input parameters depending on their probability density functions [104]. Uncertainty analysis may be used to assess the variability (prediction imprecision) in the outcome variable that is due to the uncertainty in estimating the values of the input parameters. Sensitivity analysis on the other hand extends the uncertainty analysis by identifying which input parameters are important (due to their uncertainty) in contributing to the prediction imprecision of the outcome variable. In our current model, we have adopted the Latin Hypercube Sampling (LHS) procedure developed by McKay et al [129] and has previously been applied to other epidemiological models [17, 125, 18, 126]. This method has the advantage of varying all the uncertain parameters simultaneously and the entire k-dimensional parameter space is explored. The procedure has been explained in [17, 126, 129] but in this section, we will lay out the steps that we implemented:

1. We began by identifying parameters that should be part of the analysis. To do this, we identified parameters that reported great variation between different studies as presented in the review done in Chapter 3. For parameters

whose entire parameter space was known because they lie within a well defined bounded region e.g. between 0 and 1, the entire parameter space was explored. See Table 6.4 for a list of the parameters included in the U&S analysis and their upper and lower bounds.

2. For each of the parameters we identified, we defined a probability density function from which we draw random samples.
3. We then determined the number of simulations required. There does not exist an exact formula but the following inequality has to be satisfied $N_S > \frac{4}{3}K$ where N_S is the number of simulations and K is the number of parameters involved in the analysis [17, 129]. The upper bound of the number of simulations is dependent on the availability of a computing resource that should do the work within a reasonable amount of time. We settled for 200 simulations which satisfy the inequality and was within our computing resource's ability to finish the work within a reasonable amount of time.
4. The range of each parameter was then divided into N_S non-overlapping equiprobable intervals such that $\frac{1}{N_S} = \int f(x) dx = F(x_{max}^i) - F(x_{min}^i)$ where F is the cumulative density function of x .
5. The final step involves calculation of the LHS table. This involves random sampling from each of the equiprobable regions without replacement. This ensures that a sample is taken from each of the regions forming what is known as a Latin Hypercube Sample. The sampled parameters are then paired randomly to form the input vector for the simulations.

Figure 6.2 shows a summary of the steps explained above. A sensitivity analysis

is then done using partial rank correlation coefficient (PRCC) for each input parameter and outcome variable [104]. PRCC indicates the degree of monotonicity between a specific input variable and a particular outcome variable. This method allows for the determination of the independent effects of each variable while adjusting for the variation brought about by the rest of the parameter values. Table 6.4 shows the model input parameters that were selected for inclusion in the sensitivity analysis, their upper and lower limits and the probability distribution that was adopted.

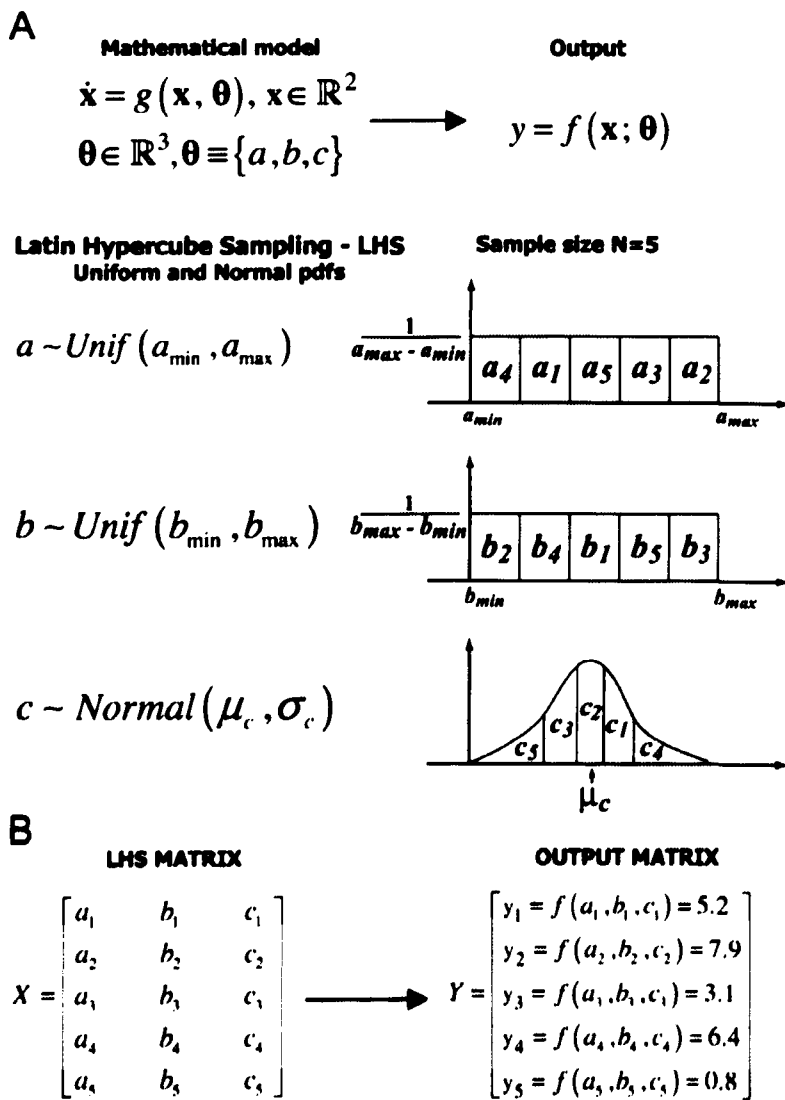


Figure 6.2: Shows the steps involved in generating the Latin hypercube sample. Adapted from Marino et al [126].

Table 6.4: Model parameters that have been included in the sensitivity analysis, their upper and lower limits and the probability function assumed

Parameter	Symbol	Lower and upper limits	Probability function
Duration of primary infection	γ_0	[4-10] days	Triangular distribution with peak at 9 days
Duration of second and subsequent infections	γ_1, γ_2	[1-5] days	Triangular distribution with peak at 4 days
Duration of short term protection	ρ_0, ρ_1, ρ_2	[2-17] months	Triangular distribution with peak at 6 months
Reduction in susceptibility after first infection	σ_1	[0-1]	Triangular distribution with peak at 0.75
Reduction in susceptibility after second infection	σ_2	[0-1]	Triangular distribution with peak at 0.65
Reduction in infectiousness of second and subsequent infections	α_1, α_2	[0-1]	Uniform distribution

6.2.5 Numerical techniques and presentation of the results

The resulting system of ordinary differential equations, shown in Eqn (6.1) is solved numerically in Matlab[®] [128] using the function *ode45* that is based on an explicit Runge-Kutta method of order (4,5) using an adaptive time step. See Appendix E for the Matlab code that we used.

A major objective in this work is to assess the effectiveness of introducing an RSV vaccine in a vaccine naive population on the incidence of total RSV hospitalizations. Predicted changes in the number of hospitalizations alone will sometimes give misleading results given that, at a single time point, a better vaccination strategy will give equivalent incidence estimates compared to incidence when there is no vaccination. This observation is due to seasonal forcing or damped oscillations before the system achieves equilibrium or its stable limit cycle. In view of this challenge, the relative effectiveness of a vaccination strategy against the baseline (no vaccination) will be measured by a case ratio (CR) defined as the ratio of the accumulated cases of RSV hospitalizations for a given vaccination strategy against the accumulated cases of RSV hospitalizations at baseline (no vaccination) [147]:

$$CR = \frac{\sum_{i=1}^{N_t} \sum_{a=1}^{N_a} H_{a,i}^V}{\sum_{i=1}^{N_t} \sum_{a=1}^{N_a} H_{a,i}^B} \quad (6.5)$$

where a is the age class, N_a is the final age class at which hospitalizations are observed, i is time in months, N_t is the final month, $H_{a,i}^B$ represents the incidence of hospitalizations under baseline scenario for each age class a at time i and $H_{a,i}^V$ represents the incidence of RSV hospitalizations under vaccination. Hence CR is < 1 if vaccination reduces the number of hospitalizations due to RSV and > 1 if vaccination increases the number of hospitalizations. This case ratio is calculated for

the models parameterized with both the diary contacts and the synthetic mixing data. To calculate the average age of primary infections at time t , denoted as $A_p(t)$, we have used the method previously used by Pitzer et al [163] and expressed mathematically as:

$$A_p(t) = \frac{\sum_{i=1}^n \psi_i \lambda_i(t) S_{0,i}(t)}{\sum_{i=1}^n \lambda_i(t) S_{0,i}(t)} \quad (6.6)$$

where ψ_i is the midpoint of the age group i , $\lambda_i(t)$ is the force of infection at time t and age i and $S_{0,i}(t)$ is the number of primary susceptibles time t and age i .

6.3 Results

6.3.1 Pre-vaccination and fitting results

6.3.1.1 Diary mixing model

In this section, we will present the model fitting results using the diary contact data. Figure 6.3 (A) shows the model fit to the age-specific profile of RSV related hospitalizations from Kilifi District Hospital while Figure 6.3 (B) shows the model fit to the time series hospitalizations. The red dots represent the hospitalization data and blue line represents the model fit. There is a good agreement between the predicted and the observed number of RSV related hospitalizations which suggests that the model is appropriate for modelling RSV transmission dynamics. Table 6.5 shows the parameters that were fitted and their optimal values together with the 95% CI (confidence interval). Figure C.1 in Appendix C confirms that the optimized values are truly locally minimum. The number of M sub-classes was only serially varied and therefore has got no CI calculated. Figure 6.4 shows the distribution of RSV cases that are due to primary (red), secondary (yellow) and

tertiary (blue) infections. It is worth noting that over 82% of the hospitalizations in children less than 13 months old are attributable to primary infection while over 80% of the infections occurring in children between 3 and 5 years of age are attributable to second infections or greater.

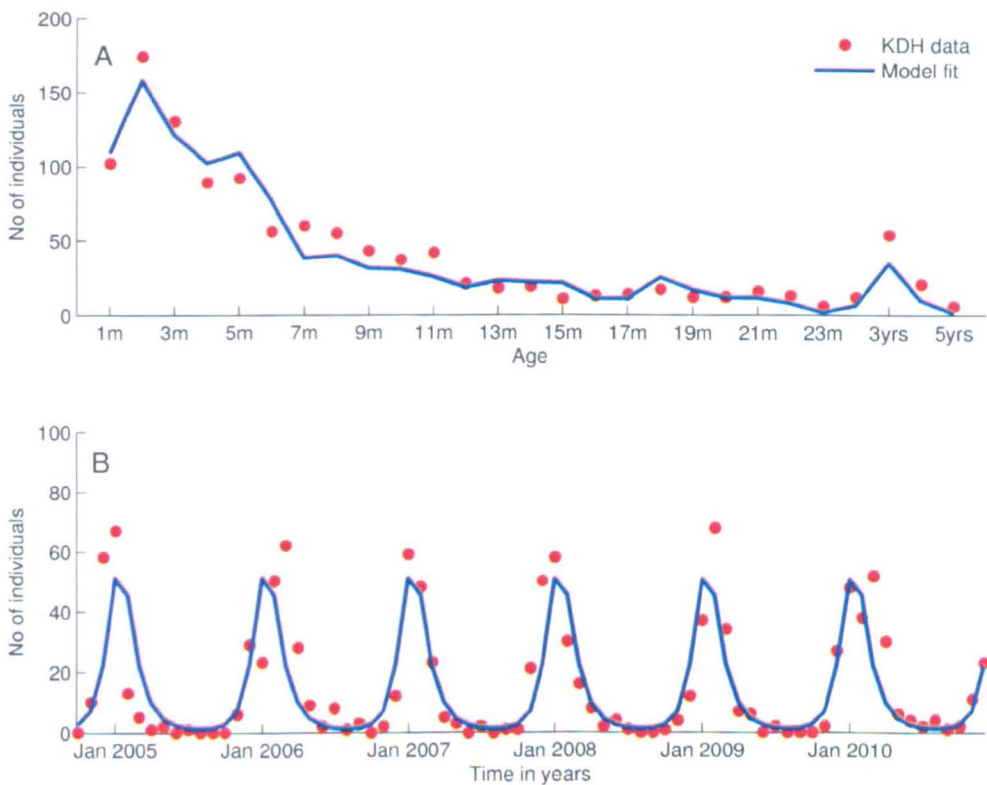


Figure 6.3: Diary model fit to age-specific (A) and time series (B) RSV related hospitalizations from Kilifi District Hospital.

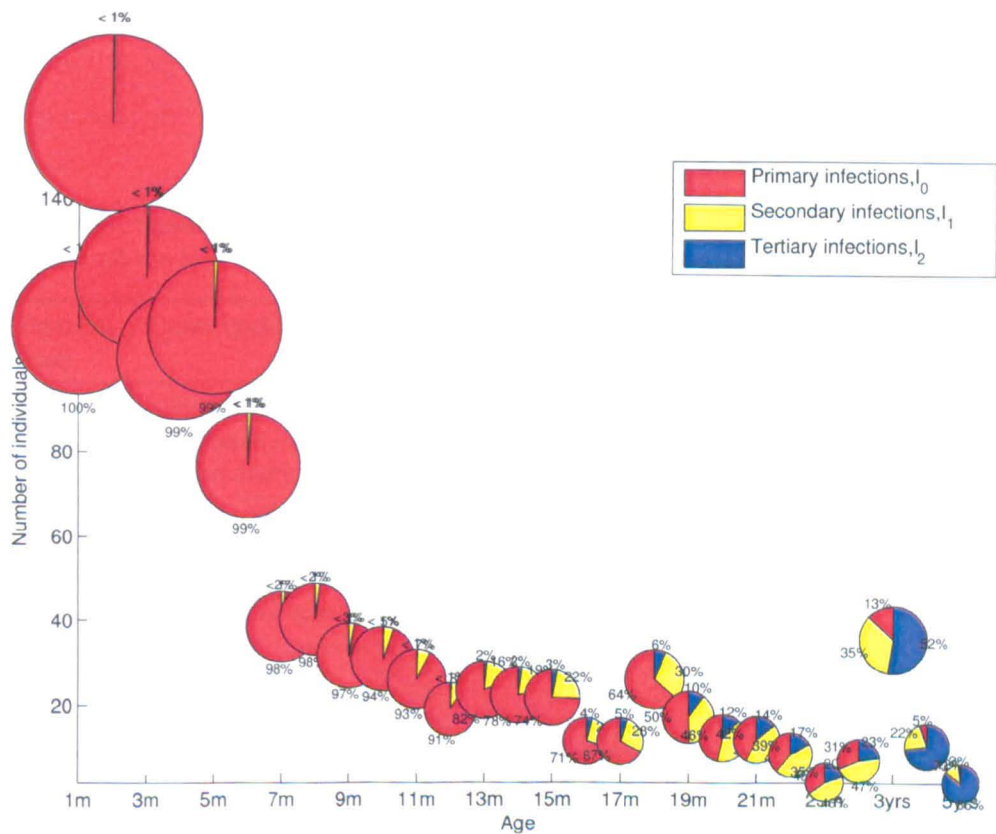


Figure 6.4: Bubble plot showing the distribution of hospital RSV related cases in each age class from the diary model fit that are due to primary (red), secondary (yellow) and tertiary & subsequent (blue) infections.

Table 6.5: Shows the parameters that were included in the fitting, their optimal values and the lower and upper 95% confidence limits for the diary contacts model

Param- eter	Description	Optimal value	Lower CI	Upper CI
q	Infectivity parameter - q value for the diary matrix	0.000679	0.000673	0.000686
a	Amplitude	0.073	0.069	0.077
ϕ	Phase angle	0.16	0.146	0.173
ω	Duration of matAb protection (months)	1.37	1.07	1.67
p	Number of maternal sub-classes	1	-	-

6.3.1.2 Synthetic mixing model

Figure 6.5 (A) shows the model fit using synthetic contact data to the age-specific profile of RSV related hospitalizations from Kilifi District Hospital while Figure 6.5 (B) shows the model fit to the time series hospitalization data. The red scatter and the blue line represent RSV hospitalization data from the KDH and the model fit respectively. From Figure 6.5 (A), it can be seen that the model tends to explain the hospitalization data well particularly from the 7th month of life, but that the 2nd, 4th and 5th months of life are least well predicted. However, from Figures 6.3 and 6.5 the two models fits are almost visually indistinguishable from each other in terms of fitting to the hospitalization data. Table 6.6 shows the parameters that we fitted, their optimal value and the 95% CI while Figure C.2 in Appendix C confirms that the optimized values are truly locally minimum. Figure 6.5 shows the distribution of the predicted hospitalized cases that are due to primary (red), secondary (yellow) and tertiary (blue) infections. Similar to the diary model, primary infections account for the majority of the hospitalizations ($\geq 79\%$) observed in children less than 13 months old while second and tertiary infections account for over 84% of hospitalizations in children between 3 and 5 years of age. By the fifth year of life, tertiary and subsequent infections account for about 86% of the hospitalizations albeit the total number of hospitalizations at this age class are quite low.

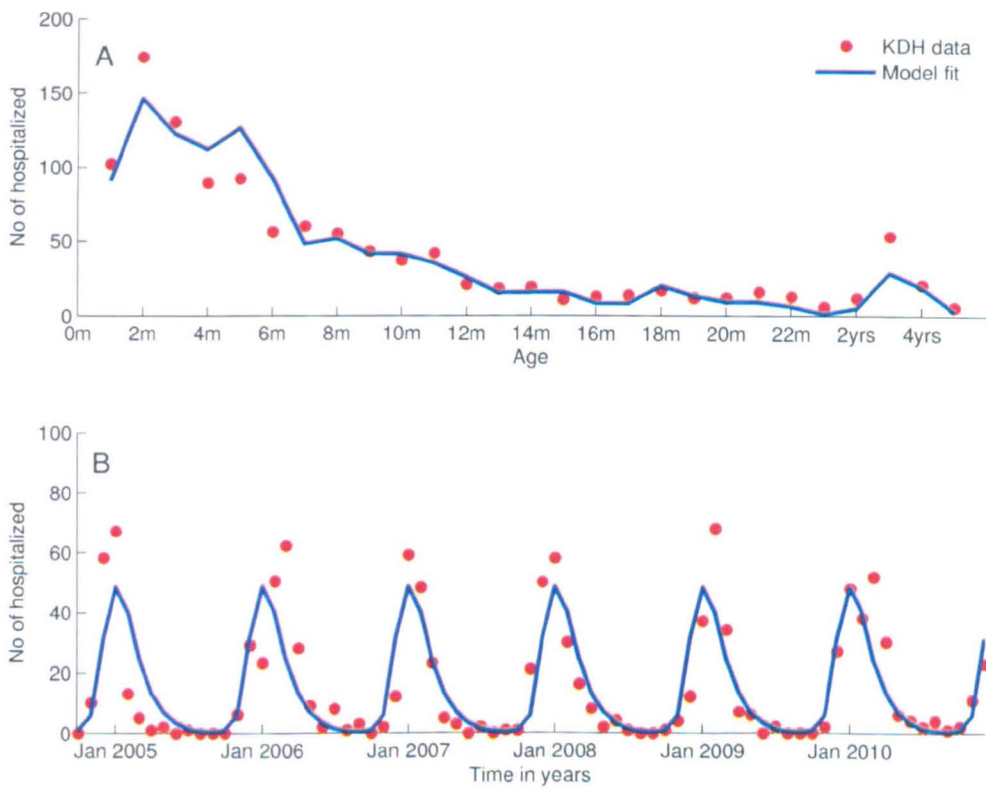


Figure 6.5: Synthetic model fit to age-specific (A) and time series (B) RSV related hospitalizations from Kilifi District Hospital.

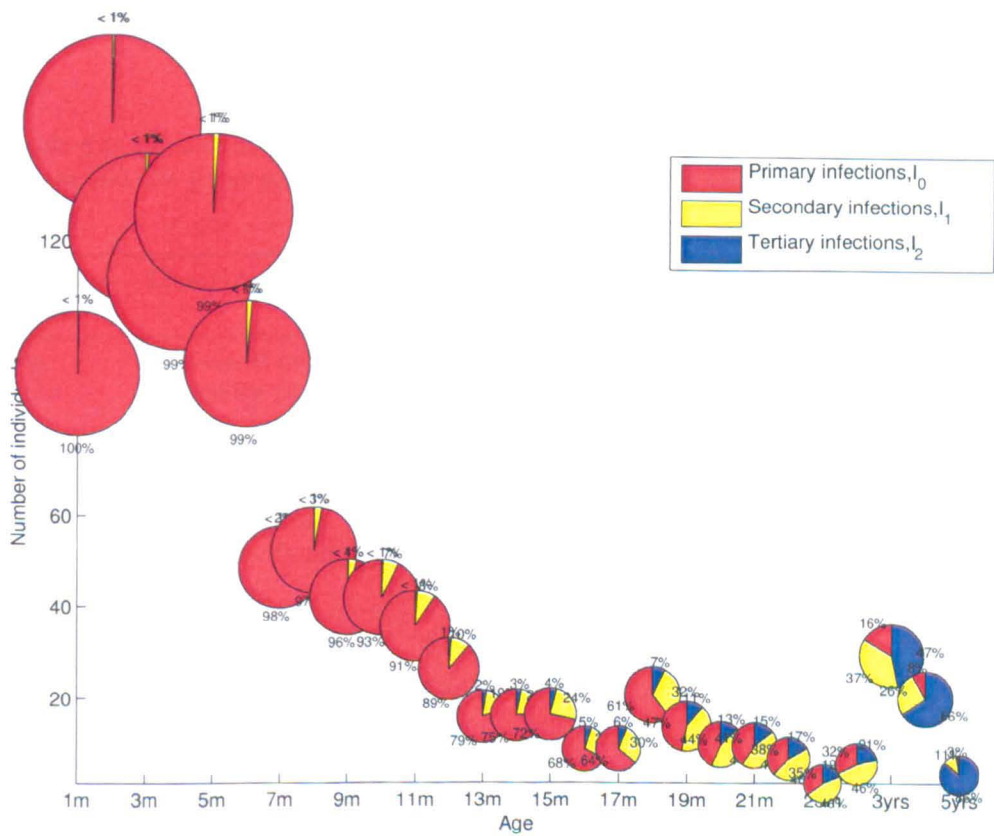


Figure 6.6: Bubble plot showing the distribution of hospital RSV related cases in each age class from the synthetic model fit that are due to primary (red), secondary (yellow) and tertiary & subsequent infections (blue).

Table 6.6: Shows the parameters that were included in the fitting, their optimal values and the lower and upper 95% confidence limits for the synthetic contacts model

Param- eter	Description	Optimal value	Lower CI	Upper CI
q_H	Household mixing infectivity parameter	96.95	95.67	98.24
q_S	School mixing infectivity factor	13,867.9	13,729.6	14,006.4
q_{HS}	General mixing infectivity factor	2.842×10^{-12}	-0.0118	0.0118
a	Amplitude	0.215	0.206	0.225
ϕ	Phase angle	2.216	2.209	2.223
ω	Duration of matAb protection (months)	3.63	3.59	3.68
p	Number of maternal sub-classes	1	-	-

6.3.2 Vaccination results

We introduce vaccination at the stable limit cycle and vaccinate both all susceptibles (all green lines in Figure 6.1) and primary susceptibles (green solid line). We begin by presenting the results when we vaccinate all susceptibles.

6.3.2.1 Vaccination of all susceptibles

Figure 6.7 shows the the proportion of hospitalizations prevented, $1 - CR$, after introducing routine vaccination at different coverages (proportion immunized) by age with (A) and (B) showing the results from the diary model and the synthetic model respectively. From Figure 6.7, we can see that the diary model leads to a greater reduction in the proportion of hospitalizations prevented compared to the synthetic model. This greater reduction in the diary model can be attributed to greater indirect benefit from vaccination as can be shown in Figure 6.11. We have then identified the *optimal age* at vaccination which is defined as the age at which the greatest vaccination benefit is achieved at the lowest vaccination coverage. Vaccination coverage is a proxy measure of the amount of resources required to reach a certain coverage i.e. the higher the vaccination coverage, the higher the resources required. The optimal month at vaccination for the synthetic model is 5 months compared to the diary model which is 7 months. For both scenarios in Figure 6.7, it is worth noting that the benefit from vaccination is increased as we delay vaccination from the first month of life to the optimal age at vaccination while a further delay decreases the impact of vaccination. However, there is little benefit lost by delaying vaccination between 5 and 13 months for the diary model. For both the mixing assumptions and even at an optimistic

vaccination coverage of 100%, all the ages at which the vaccine is evaluated do not eliminate neither the infection nor the disease. Figure 6.8 shows how the

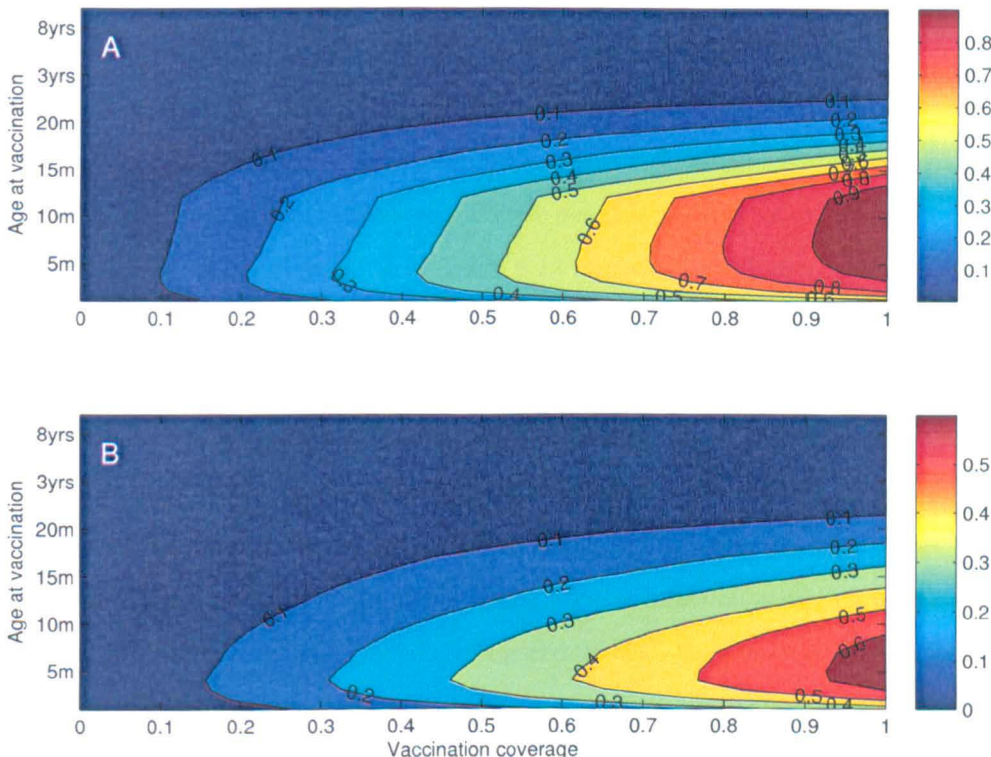


Figure 6.7: Proportion of hospitalizations prevented after introducing immunization in all susceptibles i.e. S_0, S_1 and S_2 . (A) shows vaccination outcome using the diary model while (B) is the outcome using the synthetic model. The x-axis shows the vaccination coverage as a proportion and the y-axis represents the age at vaccination.

profile of susceptible individuals at equilibrium changes with vaccination in the S_0, S_1 and S_2 classes for both the diary (A) and the synthetic model (B). Figure 6.8 (A) shows the distribution of susceptibles with the dashed and dotted lines representing the profile of susceptibles in the presence of vaccination at 7 and 15

months respectively and at 70% vaccination coverage. From the Figure, we can observe that vaccination at the optimal month i.e. at 7 months leads to the greatest reduction in the number of primary susceptible individuals. Reducing the number of primary susceptibles reduces the population which is most at risk of developing disease consequently reducing the number of hospitalizations. Vaccination also seems to increase the number of primary susceptible individuals before the month at which vaccination is implemented. This is possible because of the increase in the average age at vaccination, see Figure 6.13 implying that individuals will be susceptible to primary infection longer than before vaccination. Figure 6.8 (B) shows the profile of susceptibles for the synthetic model with vaccination at 5 (dashed line) and 15 months (dotted line) at 70% vaccination. From (B) was can observe that vaccination at the optimal month reduces the number of primary susceptible individuals but although not as much as the diary model does at 7 months. This may be as a result of the higher force of infection acting on primary susceptibles in the synthetic matrix compared to the diary matrix model, see Figure 6.10 (A).

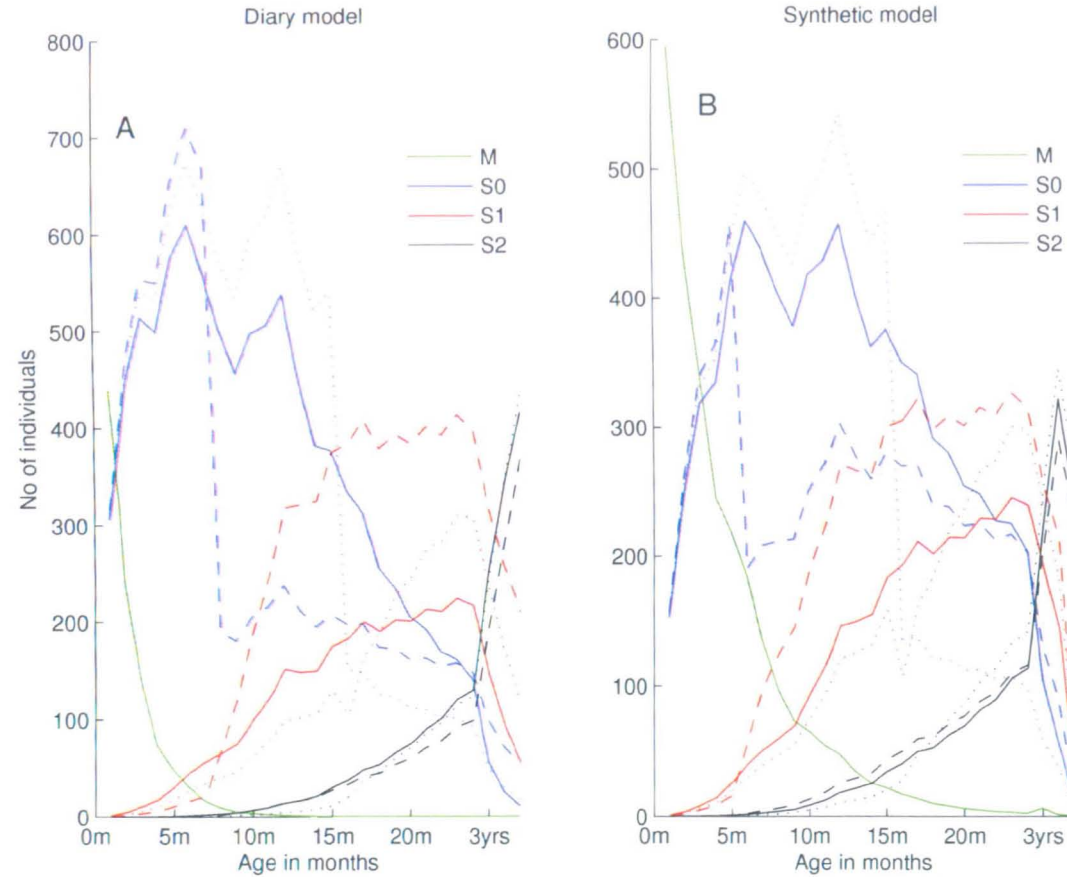


Figure 6.8: Shows the distribution of susceptibles at equilibrium for the diary (A) and the synthetic model (B). The solid lines represent the distribution in the absence of vaccination. For figure (A) the dashed and the dotted lines represent the profile with vaccination at 7 and 15 months respectively while for (B) it represents vaccination at 5 and 15 months respectively. Vaccination was implemented at 70% coverage. The green line represents the distribution of maternally protected individuals at equilibrium without vaccination.

Figure 6.9 shows the age distribution of infecteds i.e. I_0 , I_1 and I_2 at equilibrium in the presence of vaccination. Vaccination at 7 and 5 months for the diary and the synthetic model respectively (dashed lines in Figure 6.9 (A) and (B)) leads to the greatest decrease in the number of primary infections with a much greater decrease from the diary model. Indirect benefit of vaccination leads to a decline in the hospitalizations in the ages classes prior to the age at which vaccination is implemented.

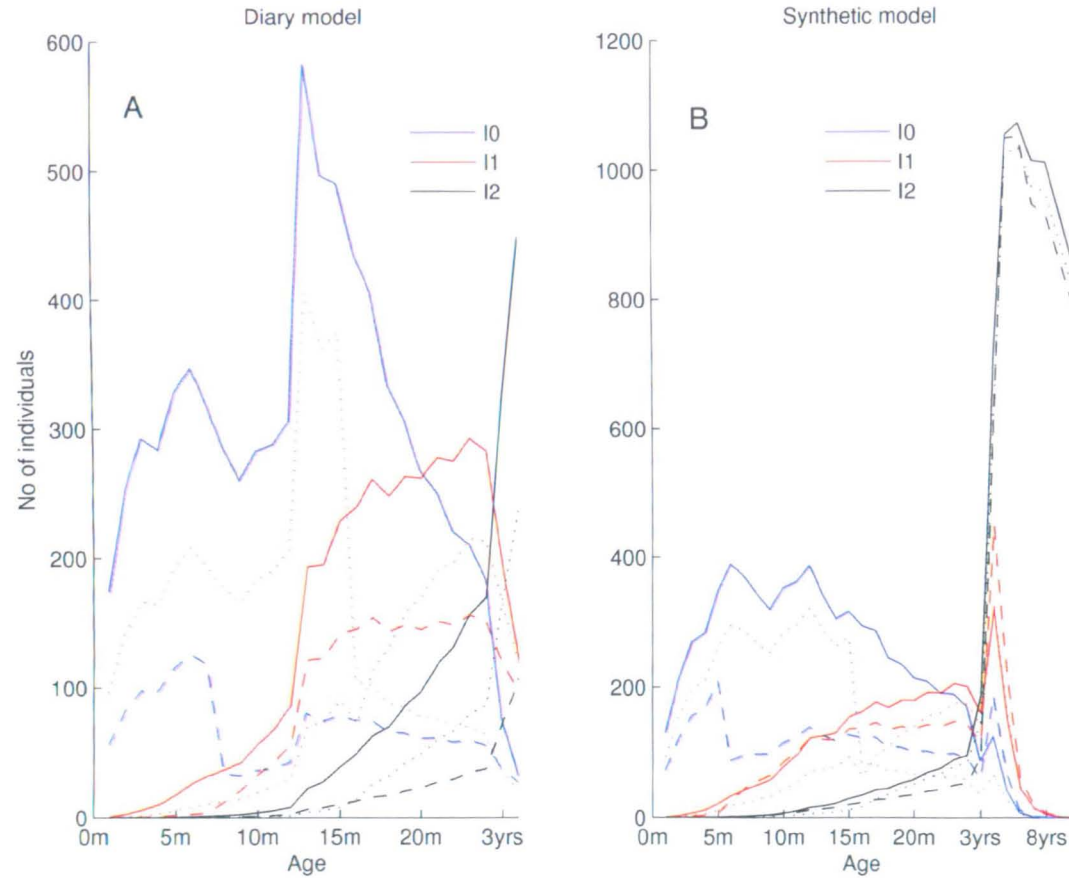


Figure 6.9: Shows the distribution of infecteds at equilibrium for the diary (A) and the synthetic model (B). The solid lines represent the distribution in the absence of vaccination. For figure (A) the dashed and the dotted lines represent the profile with vaccination at 7 and 15 months respectively while for (B) it represents vaccination at 5 and 15 months respectively. Vaccination was implemented at 70% coverage.

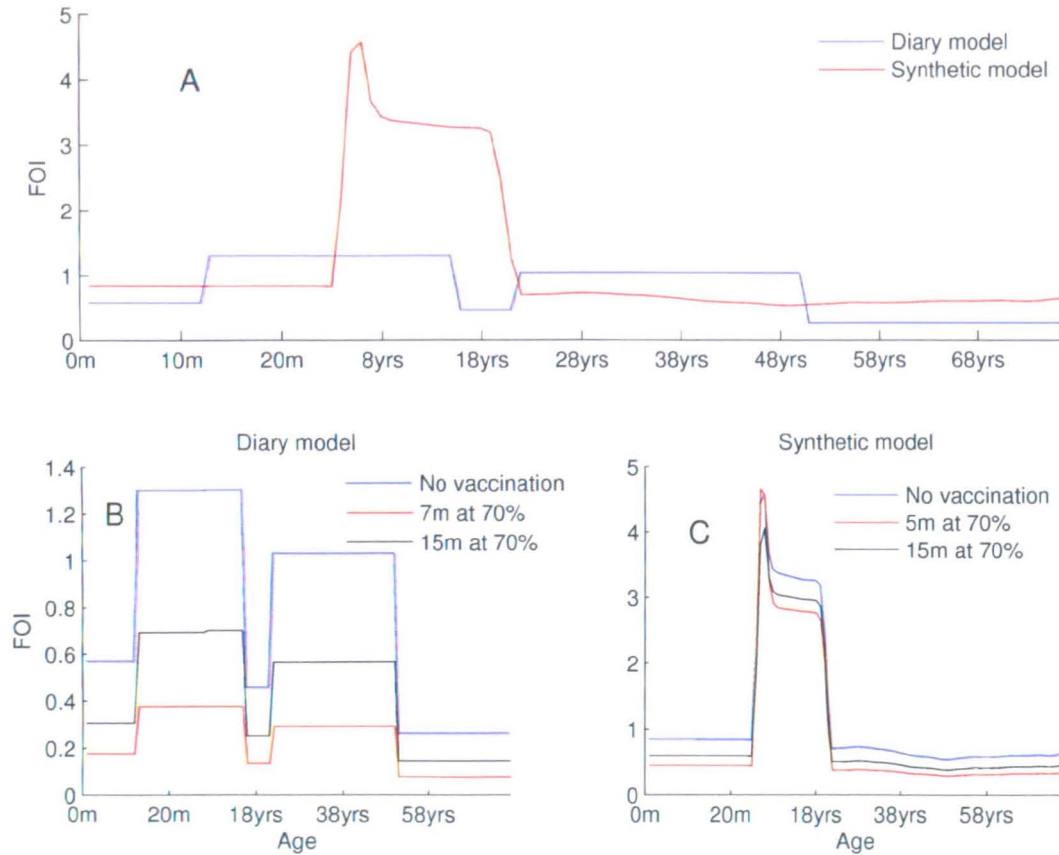


Figure 6.10: (A) shows the distribution of force of infection at equilibrium for the diary (blue) and the synthetic model (red) in the absence of vaccination. (B) Shows the profile of the force of infection for the diary model with no vaccination and with vaccination at 7 and 15 months at 70% coverage. (C) Shows the profile of the force of infection for the synthetic model with no vaccination and with vaccination at 5 and 15 months at 70% coverage.

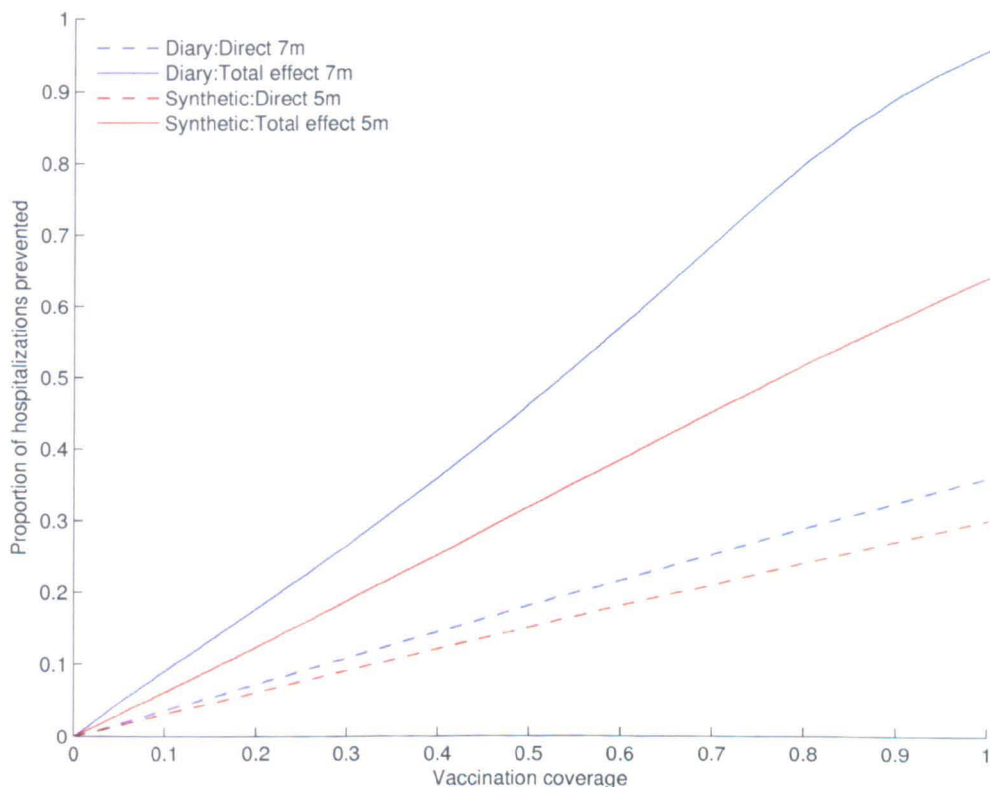


Figure 6.11: Proportion reduction in the number of hospitalizations at the optimal months for both the diary model (blue) and the synthetic matrix model (red). The dashed lines represent the direct effect of vaccination while the solid lines represent the total effect model.

Figure 6.12 shows the short-term (5 years) temporal dynamics of RSV hospitalizations after the introduction of vaccination at 7 months for the diary model (A) and 5 month for the synthetic model (B) with vaccination implemented at 70% coverage for both cases. The diary model predicts a honeymoon period of approximately 1 year when there are very few number of hospitalizations reported. The pattern of hospitalizations then results to yearly epidemics with an alternating

pattern of high and low peaks before settling down to a uniform yearly epidemic pattern. On the other hand, the synthetic model does not predict a change in the pattern except that the epidemic peaks are slightly positively skewed. In both the model assumptions, vaccination leads to a decrease in the height of the epidemics with a greater reduction observed in the diary model.

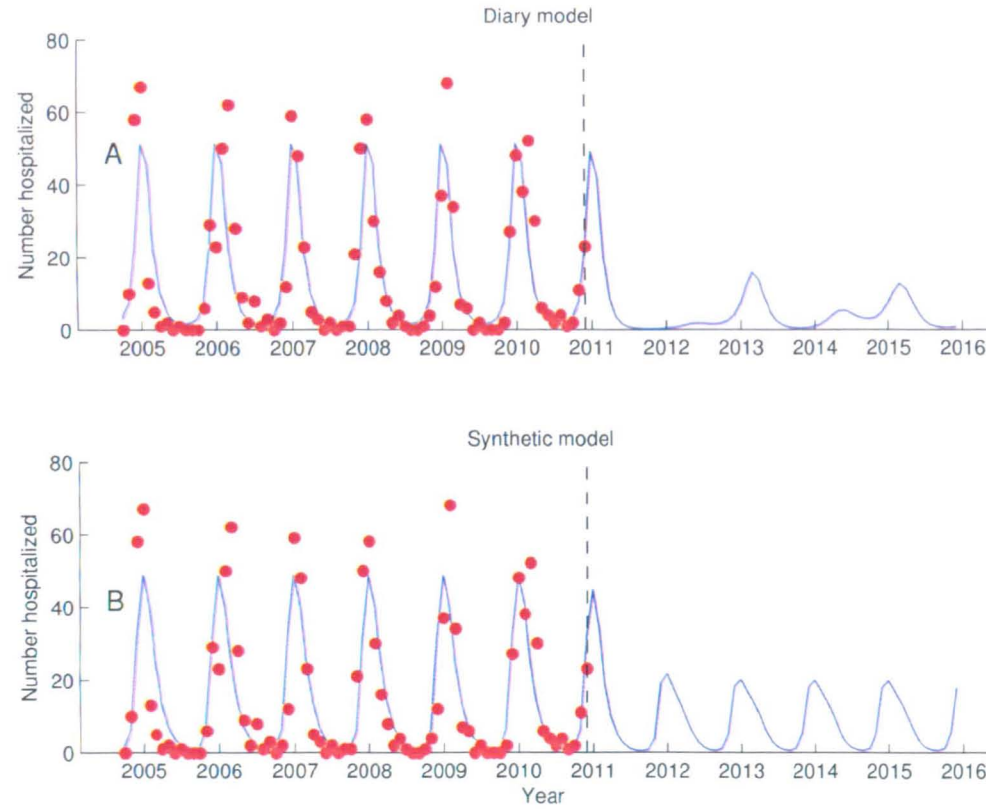


Figure 6.12: Shows the short-term (5 years) temporal dynamics of hospitalizations after the introduction of vaccination. Vaccination in the diary model (A) and synthetic model (B) is implemented at 7 and 5 months respectively at 70% vaccination coverage. The black dashed lines show the point at which vaccination is introduced.

Figure 6.13 shows how the average age at primary infection, denoted by A_p , changes as a function of both the age at vaccination and the vaccination coverage for the dairy model. In the absence of vaccination, the average age at primary infection is approximately 1.36 years. Vaccination increases the average age at primary infection with the highest value ($A_p \approx 8.33$ years) being recorded at 5 months at 100% vaccination coverage. Vaccination at the optimal month of 7 months at 100% vaccination coverage yields an average age at infection of 6.34 years. Figure

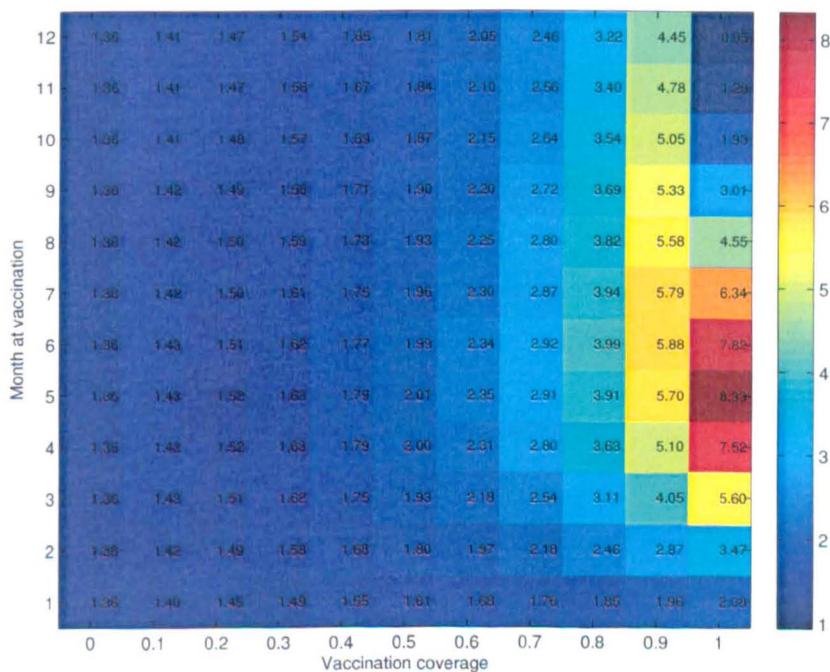


Figure 6.13: Shows the average age at primary infection (in years) as a function of both the vaccination coverage and the age at vaccination

6.14 shows the average age at primary infection for the synthetic matrix model presented as a function of both the age at vaccination and the vaccination coverage. The synthetic model's average age at infection in the absence of vaccination

is higher at approximately 1.7 years compared to the diary model's at 1.3 years. This is possibly due to the higher force of infection acting on susceptibles in the first year of life compared to the diary matrix model. Additionally, the highest age at primary infection ($A_p \approx 2.89$) is achieved at 5 months at 100% vaccination coverage. The diary model predicts an average age at primary infection which is three orders of magnitude higher than that of the synthetic model at their optimal months and vaccination coverage.

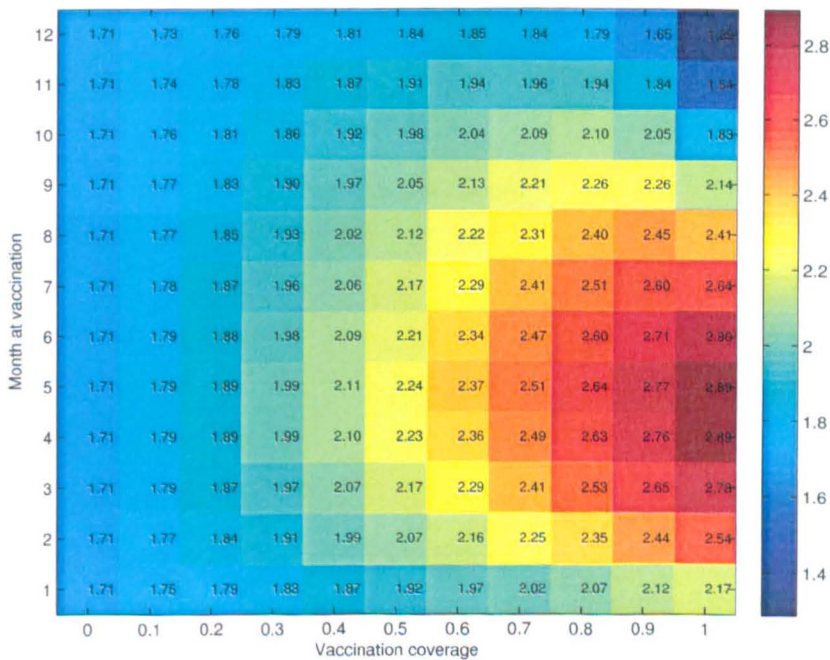


Figure 6.14: Shows the average age at primary infection (in years) as a function of both the vaccination coverage and the age at vaccination

6.3.2.2 Vaccination of primary susceptibles

In this section, we will present the impact of vaccination after introducing immunization in the primary infecteds only i.e. the green solid line in Figure 6.1. From

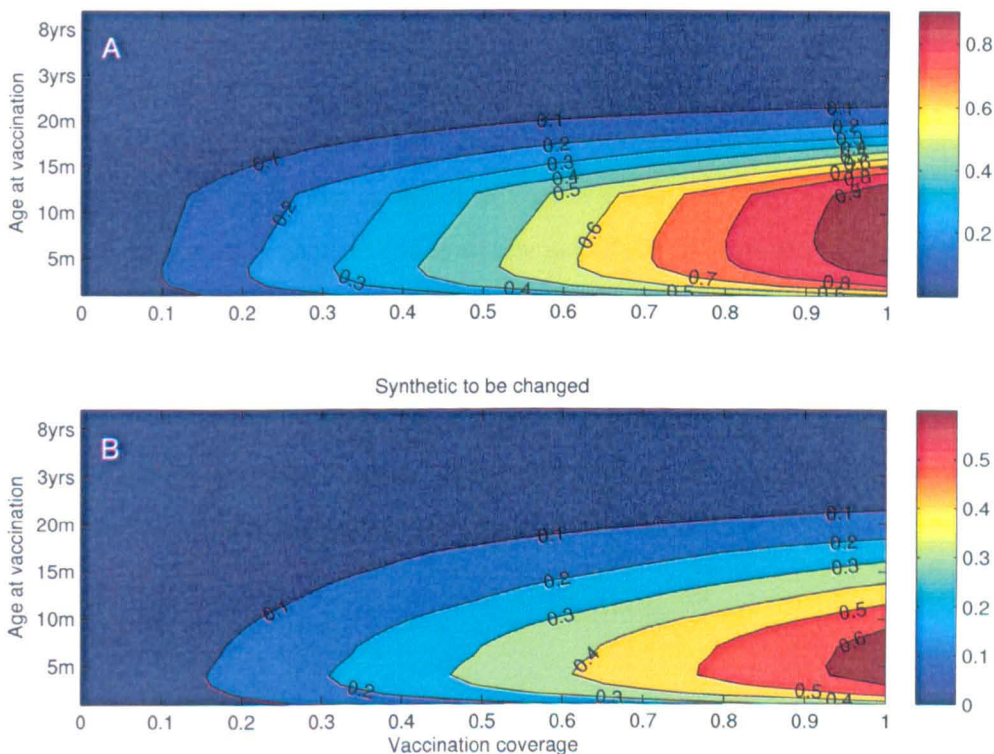


Figure 6.15: Proportion of hospitalizations prevented after introducing immunization in primary susceptibles i.e. S_0 . (A) shows vaccination outcome using the diary model while (B) is the outcome using the synthetic model. The x-axis shows the vaccination coverage as a proportion and the y-axis represents the age at vaccination. Note that there is no difference between vaccination of all susceptibles and primary susceptibles.

Figure 6.15 it is worth noting that the predicted reduction in the proportion of hospitalizations resulting from immunizing primary susceptibles is similar to that

predicted by immunizing all the susceptibles in Figure 6.7. The reason for this observation is because majority of the infections leading to hospitalizations are as a result of primary infections as can be seen in Figures 6.4 and 6.6. Over 70% of the hospitalizations, from both the diary and the synthetic model assumptions, occur in children between the age of 1 month to 15 months. A vaccine that therefore protects individuals from primary infections will achieve the greatest reduction in the number of hospitalizations. This is further supported by the age-specific profile of susceptibles at equilibrium as shown in Figures 6.8 (A) and (B) with no vaccination. Both the diary (A) and the synthetic model (B) indicate that majority of the susceptibles at equilibrium between 0 and 20 months of life are in the primary susceptible class, S_0 , and therefore vaccinating these primary susceptibles will result in the greatest reduction in the number of hospitalizations. Figures 6.9 (A) and (B) show the resulting incidence of infection from the three susceptible classes in the absence of vaccination for the diary and the synthetic model assumptions respectively.

6.3.3 Uncertainty and sensitivity analysis

The LHS technique was used to explore the effects of the uncertainty in estimating the values of the input variables on the prediction precision of two outcome variables namely: the cumulative number of hospitalizations before and after vaccination and the optimal age at vaccination. We will begin by presenting the U&S results of the diary mixing matrix and then the synthetic mixing matrix.

6.3.3.1 U&S of the diary model

Figure 6.16 (A and B) shows the imprecision in the number of hospitalization (both temporal and age-specific) that is attributable to the variation in the input parameters with the bars representing the 95% confidence intervals. The model assumptions are quite robust in predicting the number of children hospitalized with RSV since majority of the hospitalizations from the Kilifi District Hospital (red scatter plot) fall within the confidence limit. However, we observe a greater variability in the third, fourth and fifth months of life.

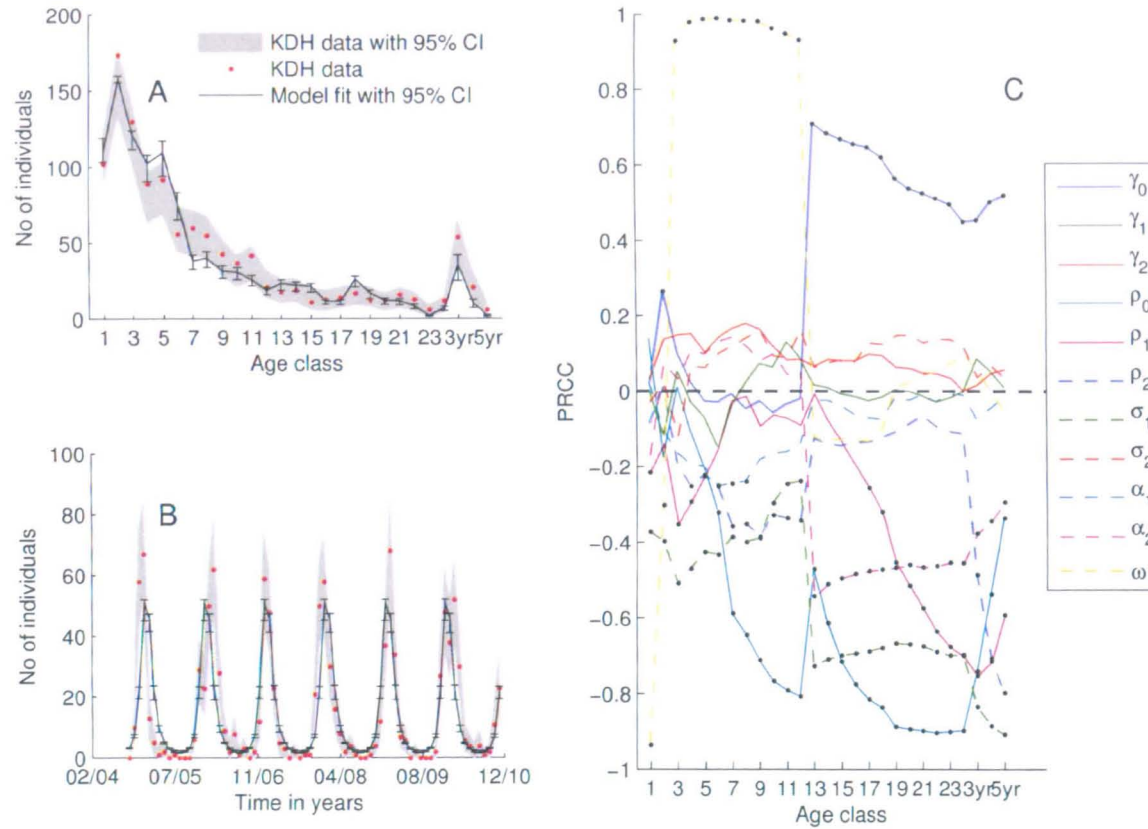


Figure 6.16: Shows the prediction imprecision in the number of age-specific (A) and time series hospitalizations (B) that is attributable to variation in the input parameters while (C) shows the PRCC per parameter for each of the age classes for the diary-based model assumption. A black dot shows that the parameter at that age class is different from zero and the difference is statistically significant.

To explore the parameters attributable to the variation observed, we present the partial rank correlation coefficient values calculated for each of the parameter for each age class as shown in Figure 6.16 (C). The sign of the PRCC identifies the qualitative relationship between the input and the output variable. A positive PRCC value implies that when the value of the input variable is increased, the number of hospitalizations (which is the outcome) increases as well and when the PRCC is negative, it implies that an increase in the input variable decreases the number of hospitalizations. A parameter with a PRCC value which is approximately zero implies that changing that parameter or even excluding it has no bearing on the outcome variable. From Figure 6.16, we can see that the duration of RSV specific maternal protection is the most important in explaining the variability in the number of children hospitalized in between the ages of 1 month and 13 months of life with PRCC values that are statistically significantly different from zero. For the first two month of life, an increase in the duration of maternal antibodies leads to a decrease in the number of hospitalizations while for children between 3 to 13 months months old, it leads to an increase in the number hospitalized. The second parameter that is important in explaining the variability observed in the number of hospitalization is the duration of primary infection i.e. γ_0 . An increase in the duration of primary infection leads to an increase in hospitalization in children older than 14 months old while for younger children children it leads to a reduction in the number of hospitalizations. Thirdly, an increase in the duration of immunity in all the three immune classes i.e. ρ_0 , ρ_1 and ρ_2 leads to a decrease in the number of hospitalizations in all age classes. This implies that an introduction of a vaccine that increases the duration of protection from infection would be beneficial in reducing the number of hospitalizations.

Figure 6.17 (A) shows the imprecision in the optimal month at vaccination that is due to the uncertainty in measuring the input variables. The optimal month at vaccination is 7 months with the 95% CI ranges from 3 to 12 months. Figure 6.17 (B) shows a tornado plot with the PRCC values on the x-axis and the input parameters on the y-axis with the red scatter besides the bar showing that the value is statistically significantly different from zero. From the figure, the parameter that is contributing to the greatest variability in the optimal month at which to vaccinate is the duration of RSV specific maternal antibodies. An increase in the duration of maternal antibodies leads to an increase in the optimal age at vaccination. An increase in the duration of infection (γ_2) and infectiousness of tertiary infections (α_2) and an increase in the susceptibility of tertiary susceptibles (σ_2) leads to a significant decrease in the optimal age at vaccination since an increase in these parameters leads to an increase in the force of infection. On the other hand, an interesting result is that an increase in the infectiousness (α_1) of second infections leads to a reduction in the optimal age at vaccination.

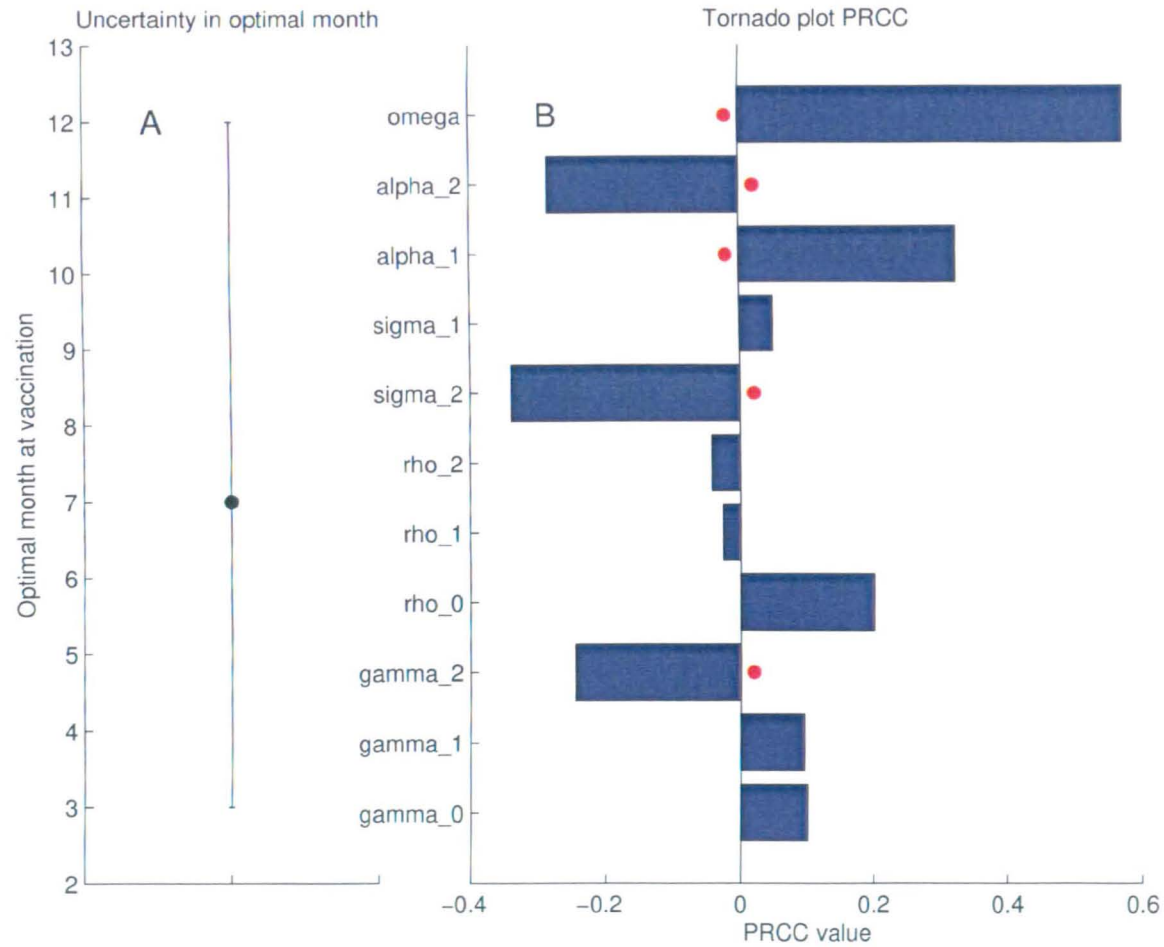


Figure 6.17: (A) shows the prediction imprecision in the optimal month at vaccination that is attributable to variation in the input parameters while (B) shows the PRCC for each parameter values for the diary model. Red scatter points beside the bars represent the parameters that are different from zero and the difference is statistically significant.

6.3.3.2 U&S of the synthetic model

Figure 6.18 (A) and (B) shows the variability in the number of hospitalizations (both age-specific and temporal) that is attributable to the variation in the input parameters with the bars representing the 95% confidence interval of the model prediction. Most of the age-specific hospitalization data is predicted to fall within the 95% confidence interval of the model fit as can be seen in Figure 6.18 (A). However, there is also a lot of variability that is observed in the first six months of life. Most of the variation observed in the first twelve months of life can be attributed to uncertainty in four main parameters 1) duration of protective effect of maternal antibodies, 2) duration of primary infection, 3) duration of protection after primary infection and 4) duration of protection after second infections as can be seen in Figure 6.18 (C) from the values of their PRCC. Increasing the duration of maternal antibody protection reduces the number of hospitalization in the first three months of life. This is because individuals are protected from infection at the age range when they are most susceptible to severe disease requiring hospitalizations. On the other hand, this increase leads to an increase in the number of hospitalizations between the 5th month and the 3rd year of life. This result can be attributed to the loss of maternal antibodies with individuals becoming susceptible to infection at an age when they are still vulnerable to severe disease. The importance of the duration of RSV specific maternal antibodies starts to decline beyond the 3rd year of life.

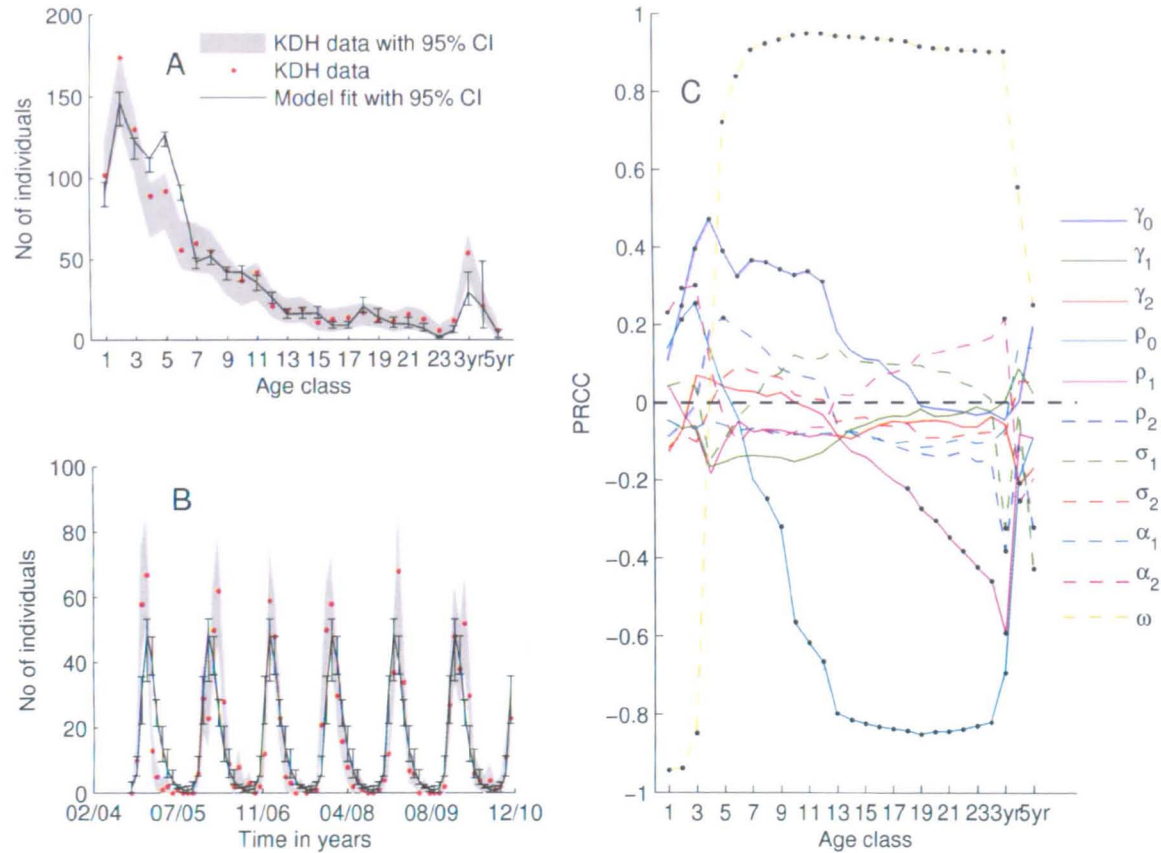


Figure 6.18: Shows the prediction imprecision in the number of age-specific (A) and time series hospitalizations (B) that is attributable to variation in the input parameters while (C) shows the PRCC per parameter for each of the age classes for the synthetic model assumption. A black dot shows that the parameter at that age class is different from zero and the difference is statistically significant.

An increase in the duration of primary infection leads to an increase in the number of individuals hospitalized between the 2nd and the 11th months of life after which it ceases to be an important parameter in explaining the variation. An increase in the other two important parameters i.e. duration of primary and secondary protection, leads to a decrease in the number of hospitalizations since increasing the duration of protection reduces the number of infecteds in the population by reducing the per capita rate of infection. Most of the other parameters have their PRCC values around zero and are therefore not important in explaining the variation.

Figure 6.19 (A) shows the imprecision in the optimal month at vaccination that can be attributed to the variation in the input parameters. The optimal month at vaccination is 5 months with the 95% CI ranging from 4 to 9 months. The tornado plot in subplot B shows the PRCC values for each of the parameters. Six parameters are important in explaining the uncertainty in the optimal month at vaccination as shown by the red scatter dots besides the bars and they are: the duration of primary and second infections (γ_0 and γ_1), level of susceptibility of third and subsequent infections (σ_2), the infectivity of second and subsequent infections (α_1 and α_2) and the duration of maternal antibodies protection (ω). Except for the duration of maternal antibodies, an increase in the rest of the parameters lead to an increase in the optimal age at vaccination. This is possibly so because an increase in these parameters leads to an increase in the number of infecteds in the population consequently increasing the force of infection hence reducing the average age at primary infection leading to a vaccination strategy that requires early delivery of the vaccine. An increase in the duration of maternal antibodies leads to an increase in the optimal age at vaccination.

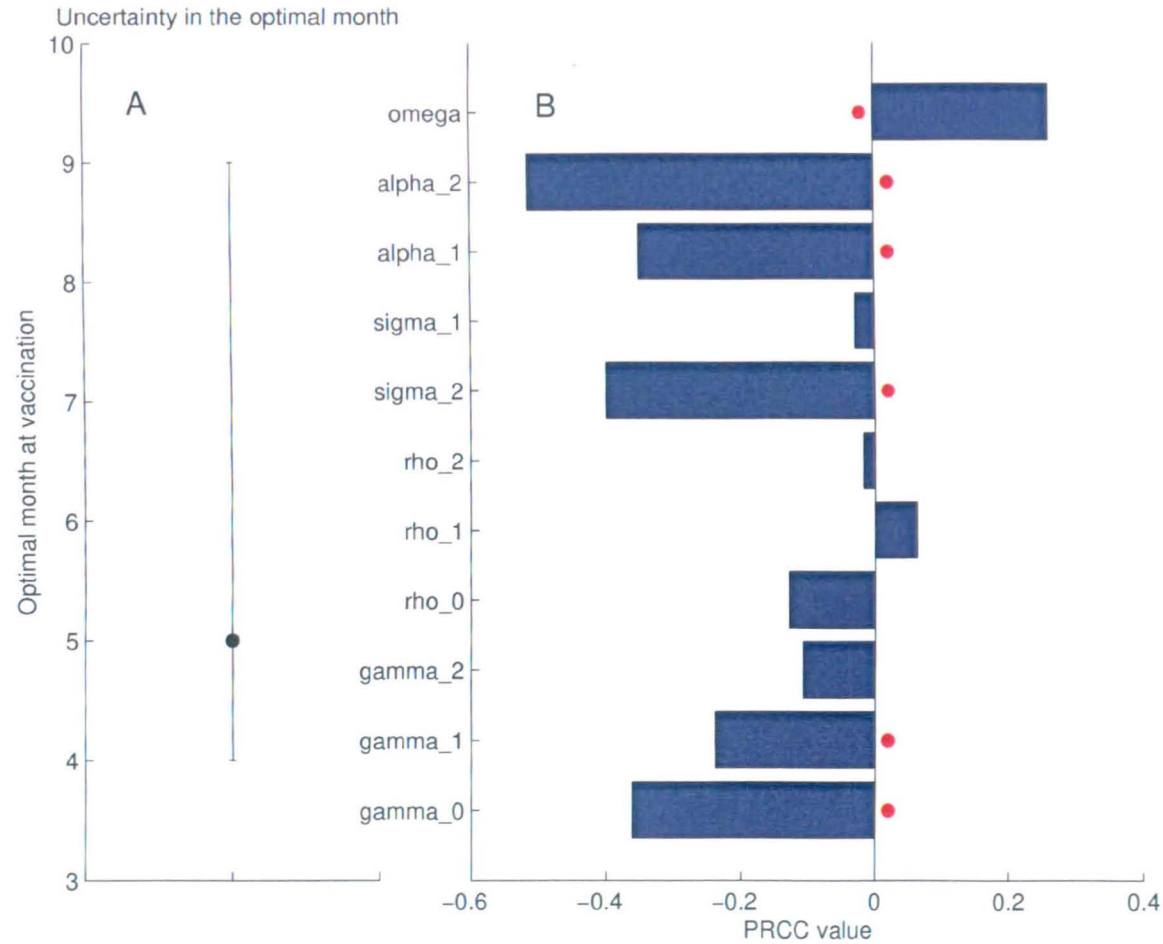


Figure 6.19: (A) shows the prediction imprecision in the optimal month at vaccination that is attributable to variation in the input parameters while (B) shows the PRCC for each parameter values for the synthetic model. Red scatter points beside the bars represent the parameters that are different from zero and the difference is statistically significant.

6.4 Discussion

In this chapter, we have presented the analysis of a mathematical model describing the transmission dynamics of RSV. Additionally, we have explored the long-term impact of introducing vaccination in the population with two mixing assumptions. Hospitalization output from the model with both mixing assumptions fits well to both the age-specific and time series hospitalization data. For the model assuming the synthetic matrix we have estimated three q values: q_H (household mixing), q_s (school mixing) and q_{HS} (general mixing) with q_{HS} giving a lower confidence limit which is below zero. The confidence interval for the q_{HS} is wide suggesting that the hospitalization data was probably too limited to provide information or support the inclusion of the homogeneous mixing matrix in the linear combination generating the WAIFW matrix. Therefore, the conclusions derived from the model with the synthetic matrix warrant to be compared with the model fitted to other hospitalization datasets.

It is quite clear that the introduction of a vaccine at any age between 1 month and 5 years leads to a reduction in the incidence of hospitalizations in children without any adverse effects. However, it has been demonstrated in this work that vaccination leads to an increase in the average age at first infection. This has previously been observed with the introduction of vaccination of other childhood infections [96, 8, 164]. The clinical effects of a first infection at an older age for RSV is not documented since majority of the primary infections occur within the first two years of life [68]. If the development of severe disease is, at least in part, physiological, then older individuals who have a greater residual for gaseous exchange, may have a reduced risk of respiratory congestion. It has been shown

that the severity of infections decreases with increased age at infection [149, 93] and by the history of infection [93]. However, if disease does not decline with age at first infection, or even increases, then the model will overestimate the reduction in the number of hospitalizations with the introduction of vaccination. Under this scenario, vaccination would increase the average age at primary infection and if getting the first infection later in life is detrimental, then vaccination would lead to more severe first infections. However, we currently do not know the clinical outcome of a delayed first infection since most children will have their primary infection in their first two to three years of life [93, 68].

Routine vaccination at 7 and 5 months are the optimal vaccination strategies in reducing the number of hospitalizations for the diary and the synthetic mixing models respectively. For the model using the diary data from Figures 6.7 (A) and 6.15 (A), there is little loss in effectiveness resultant from delay to vaccination is between 5 to 12 months of age. On the other hand, there is a sharp decline in the impact of vaccination in the synthetic model beyond the optimal vaccination month. There are two observations here: 1) we see a flat area of peak benefit between 5 months and 12 months for the diary model and a peaked profile for the synthetic model around 5 months. This result seems to be a trade-off between decay of protective RSV specific maternal antibody and the rate of infection. Although there is a predicted longer duration of maternal antibodies from the synthetic matrix, i.e. 3.6 months compared to 1.4 months for the diary model, this does not result in delayed optimal age at delivery since as soon as they lose their protection they get infected due to the high force of infection in the early age classes as can be see Figure 6.10 (A). On the other hand, the diary matrix predicts a rapid loss of maternal antibodies hence the benefit from early vaccination but

little in the way of losing the benefit with increasing age due to the relatively lower force of infection in children below 12 months. However, at 13 months, the force of infection rises and this leads to a sharp decline in the benefit from vaccination in the diary model.

Vaccination of children older than 15 months results in very little or no reduction in the number of hospitalizations implying that children less than 15 months are the ones who are important for the transmission of RSV within a community. If the protective effect of vaccination is short lived, since the vaccine behaves like a natural infection, why then do children less than 15 months seem important? This is the question that comes into mind when one sees that the highest benefit of vaccination is in children less than 15 months instead of the previously hypothesized school going children [156, 140] who bring the infection back to the household. This is partly because protecting the naïve susceptibles will directly prevent disease. As can be seen in Figures 6.4 and 6.6, majority of the hospitalizations are attributable to primary infections. However, based on these results, it would be premature to write off school child vaccination. It seems that the model structure does not support household mixing which is required in order to have strong mixing in the household that leads to transmission from school child to infant through other members of the household. The structure of the contact matrices in Figures 4.7 and 4.8 reveal this fact. In that there is high mixing within and between the school age groups but not a lot of the off diagonal mixing with infants i.e. mixing is held within the school groups and does not impact on infants. Another possible explanation why vaccination of the older individuals leads to very minimal benefit is because immunity is short lived and therefore it does not reduce the force of infection appreciably to provide indirect protection. Our understanding is

that vaccination across all age groups would however have a short term benefit deferring infant infection, but would need to be repeated annually for there to be a lasting effect.

From the uncertainty and sensitivity analysis, it is quite clear that the duration of RSV specific maternal antibody is important in explaining the variability observed in both the age-specific hospitalizations and the optimal month at vaccination for both the diary and the synthetic model assumptions. For the diary model, an increase in the duration of RSV specific maternal antibodies reduces the number of hospitalizations in the first 2 months of life while increasing the number of hospitalizations in the 3rd to the 13th months of life. For the synthetic model, the duration of maternal antibodies leads to a decrease in the number of hospitalizations in the first three months of life while increasing the number of hospitalizations between the 4th month and the 3rd year of life. Beyond this point, maternal antibodies cease to be a significant. This analysis has got important epidemiological implication. It has revealed that the model prediction of the number of hospitalizations is robust, only a few of the variables are important in explaining the variability with the duration of maternal antibody being important in both model assumptions. This suggests that it is important to quantify this variable accurately. Reducing the uncertainty increases the prediction precision of the model. Additionally, the importance of the duration of RSV specific maternal antibody reveals that vaccinating pregnant mothers would be potentially beneficial in reducing the amount of disease in early infancy if it leads to an increase in the duration and level of maternal antibody protection in the offspring. This will allow for protection from infection at an age when the child is most susceptible to severe disease and also allows for vaccination at an older age when RSV vaccines

have been shown to be safe and immunogenic. Uncertainty in the optimal month at vaccination is between 3 and 12 months for the diary-based model while that of the synthetic mixing matrix model is 4 to 9 months. Both models suggest that delaying the age at which a vaccine is delivered from birth is beneficial.

For the diary model, an interesting observation is the effect that α_1 and α_2 have on the optimal age at which the vaccine should be delivered. The expectation is that an increase in the infectiousness of both I_1 and I_2 will increase the force of infection hence increasing the the level of infection in the population leading to a reduction in the optimal month at which the vaccine should be delivered. This is true for α_2 . However, an increase in α_1 leads to a counter intuitive outcome of increasing the optimal age at vaccination. This outcome seems to be robust to the number of LHS samples since increasing the number of LHS samples from 100 to 200 does not affect the outcome. This outcome should be looked at in the light of the changes in the other parameters as well. For example, increasing α_1 generally leads to an decrease in the optimized infectivity parameter denoted as q . Therefore, increasing α_1 drives the overall dynamics that tend to decrease the transmission potential in the population if looked at in the presence of changes in the other parameters. This is contrary to what one would expect if α_1 alone is increased. The non-linear effect of the changes in the parameters operating through the different mixing matrices might result in the differences between the diary and the synthetic matrix outcomes for α_1 .

We have only implicitly accounted for the antigenic differences between strains of RSV. Earlier work demonstrated that inclusion of two genetic types, groups A and B, explained some of the epidemiological pattern, especially the group A and B epidemic dominance [206]. Secondary infections may also only occur if

a susceptible individual encounters an RSV strain that is substantially different from the previous infection as suggested by Agoti et al [3]. Consequently, the host immune responses to the two strains is likely important and may influence the outcome of vaccination. Additionally, the characteristics of a vaccine in terms of its ability to generate immunological responses across the spectrum of RSV antigenic types will greatly influence the outcome of vaccination [45, 170]. Although our model produced the temporal features of an RSV epidemic, the exact factors that drive the epidemic are not explicitly known. We have modelled the epidemics using a cosine function with annual forcing. It is interesting to note that even with this kind of a forcing function, introduction of a vaccine eventually changes the pattern of epidemics for the diary-based model. More work needs to be done to explore the potential factors driving the the epidemics.

We plan to extend this work by exploring different vaccination strategies alongside the routine vaccination of children. A number of strategies have been proposed in literature: maternal immunization in late pregnancy [50, 161], vaccination of school going children [169] and campaign vaccination over a wide age range.

Chapter 7

Modelling transmission dynamics in households

7.1 Introduction

In early models developed to describe the transmission of respiratory pathogens, the population was ordinarily assumed to mix homogeneously with frequency or density dependent transmission [116, 10]. These homogeneous models were later extended to account for host heterogeneities in the transmission. This included the stratification of mathematical models to correspond to the age structure of the host population [180, 47, 10, 116] where age was used as a measure of elevated risk e.g. increased infectivity or acquisition of infection with age. Further extension has taken a different form where the population is divided into various risk groups [117, 116] and more recently by households [14, 102, 62, 146].

For a number of respiratory infections requiring close contacts, their transmission within the household or the family is an important mechanism for their spread

[102, 209] due to the greater strength of contacts between individuals sharing living arrangements compared to contacts outside of the household [122, 166]. There are a number of other aspects of a household that make them special and epidemiologically more different than other heterogeneities. Firstly unlike schools and work places, they have a wide range of ages and gender allowing for inter-generational transmission of infection more likely to occur. Secondly, there is greater genetic similarity within household members than between random individuals. Such heterogeneities makes households epidemiologically relevant since the ease or difficulty with which an infection occurs or controlled is dependent on the factors mentioned above albeit not limited to them. For example a recent household study demonstrated that older children particularly school going are the frequent introducers of RSV into households that lead to infant infection [140]. Exploring transmission dynamics using household models can also allow for more targeted implementation of interventions. House & Keeling [102] have demonstrated that targeting to vaccinate households with more susceptibles is a better strategy compared to vaccinating random households or random individuals against influenza and hence can allow for more efficient use of vaccines.

There has been recent interest in household modelling with earlier models dealing with SIS-type model dynamics in household structured settings [13, 62, 146]. Ball [13] demonstrated that the stochastic household model can be approximated using a deterministic model. In order for the deterministic model to provide a good approximation to the more realistic stochastic model, the author suggests that it is necessary that the number of households in the population be assumed infinite so that the probabilistic effects leading to stochastic extinction can be averaged out. This has also been demonstrated by Neal [146]. Following this early

work, more formal household models making use of the deterministic approximation have appeared linking data with the models [101] with others exploring the effect of interventions [102]. More recently, Ross et al [175] have extended the methods developed earlier to include more efficient methods of evaluating some of the more important epidemiological quantities such as the invasion threshold, early growth rate, household offspring distribution and endemic prevalence of infection using path integrals for Markov Chains.

In this chapter, we will present a simple multi-strain deterministic household model describing the transmission of RSV both within a household and in the community. The model is composed of two transmission rates: one representing transmission between members of the same household and the other transmission to general members of the community. This simplified model has been used to describe the equilibrium dynamics of RSV in the general population and within the household. The motivation for presenting this work here is two fold a) we want to demonstrate that the household model allows for more epidemiologically relevant parameters to be determined compared to the model considering age only as the most important heterogeneity and b) this model would form a basic framework on which to develop new methods of analyzing household cohort data. Recently, these kind of studies are being done with increasing frequency and they seek to measure infections in whole households over time.

The deterministic household model developed in this chapter was chosen due to its simplicity of conception and the ease of parameterization. An alternative formulation would take the form of an individual based model which can be readily iterated with any of the standard stochastic algorithms e.g. the Gillespie's direct algorithm [63]. However, computational difficulties would arise due to the vast

number of different possible events. Parameterization of the model would also be difficult to implement [116]. In deterministic settings, parameters are chosen that minimize the deviation between observed and simulated epidemics so that the predicted behaviour by the system matches, as close as possible, the observed behaviour. However, stochasticity can have a significant effect on the mean and therefore possibly bias this simple form of estimation. This problem is enhanced when localized extinctions are frequent especially when implementing an integer-valued stochastic model.

7.2 Objectives

In the work presented in the chapter, we seek to develop a household based model and assess its potential in describing the transmission of RSV A and B within a population of households.

7.3 Methods

In order to achieve the objectives mentioned above, we have adopted the methodology used by House & Keeling [101] which is a deterministic model with explicit household structure. For the epidemiological model, we have adopted a modified version of the multi-strain RSV model developed by White et al [206]. Figure 7.1 shows the graphical representation of the epidemiological compartments and the flow between them. In this work, we take a simpler approach to the epidemic model so as to focus exclusively on the implications of household structure on the transmission dynamics. Briefly, individuals are initially susceptible to primary in-

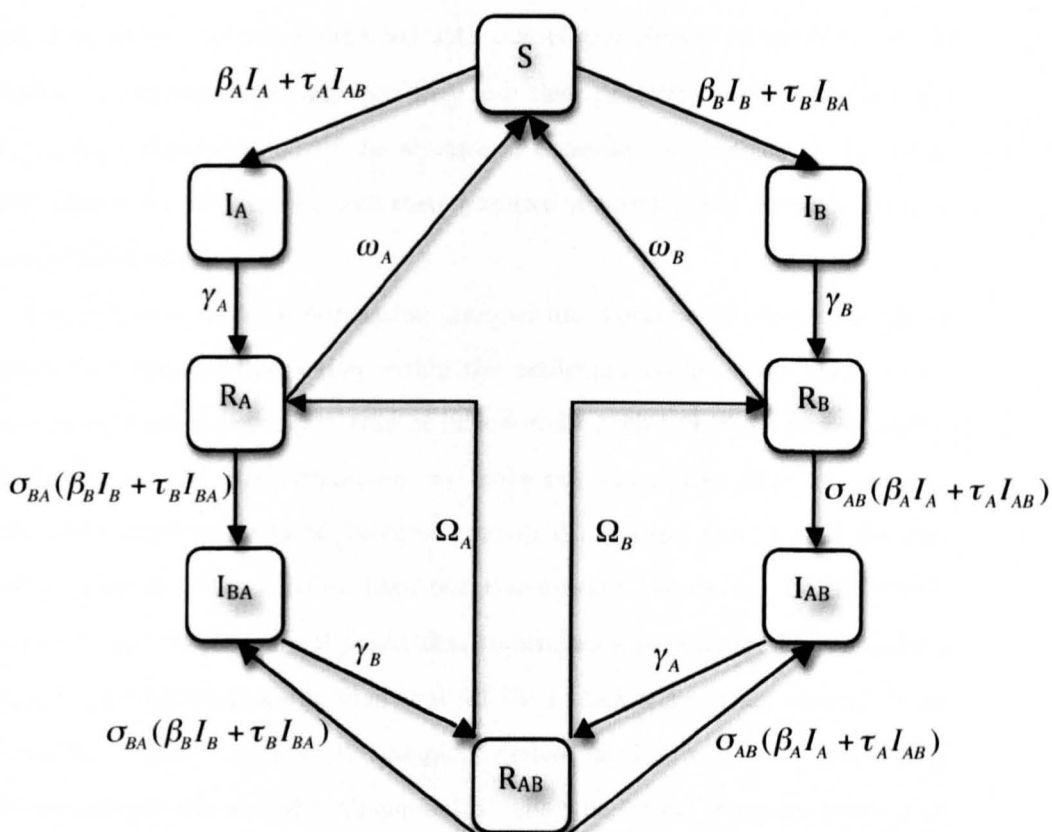


Figure 7.1: Graphical representation showing the epidemiological compartments and flow of individuals between them. See text for definition of symbols.

fection with either group A or B virus with strain specific transmission parameter. If infected, they recover into the R_A and R_B , depending on the primary infecting strain, where they remain susceptible to a second infection albeit at a lower risk of infection compared to primary susceptibles in class S . At this point, re-infection will be governed by the cross protection arising from the previous infection. Once infected with a heterologous strain, individuals recover into the R_{AB} class. This class contains individuals who have been infected with both strains. Individuals

can then either, become re-infected with one of the viruses (A or B) or in the absence of exposure to re-infection they lose their protection and revert into the R_A or R_B compartments. In the absence of exposure, individuals in the recovered classes R_A and R_B will lose their acquired immunity and move to primary susceptible class, S .

Let us now consider a population grouped into households where individuals retain their random interaction within the entire population while they experience an additional per capita rate of infection for each infectious person within the household. In this formulation, we make the assumption that transmission potential within members of the same household is greater than within the general population. We therefore have two transmission parameters 1) the within household and 2) the general population transmission parameter. Let us make a simplifying additional assumption that all the households contain exactly N individuals. Figure 7.2 shows the coupling system used to define the interaction between households and within households. The blue arrows represent interaction between households while the black arrows represent stronger interaction within households.

To write down the model equations, let us define $H_{a,b,c,d,e,f,g,h}$ as the proportion of households in the population consisting of a, b, c, d, e, f, g, h individuals in the $S, I_A, I_B, R_A, R_B, I_{BA}, I_{AB}$ and R_{AB} compartments respectively with $\sum H_{a,b,c,d,e,f,g,h} = 1$ and with the number of people represented by the alphabets from a to h are integers. To calculate, for example the proportion of susceptible

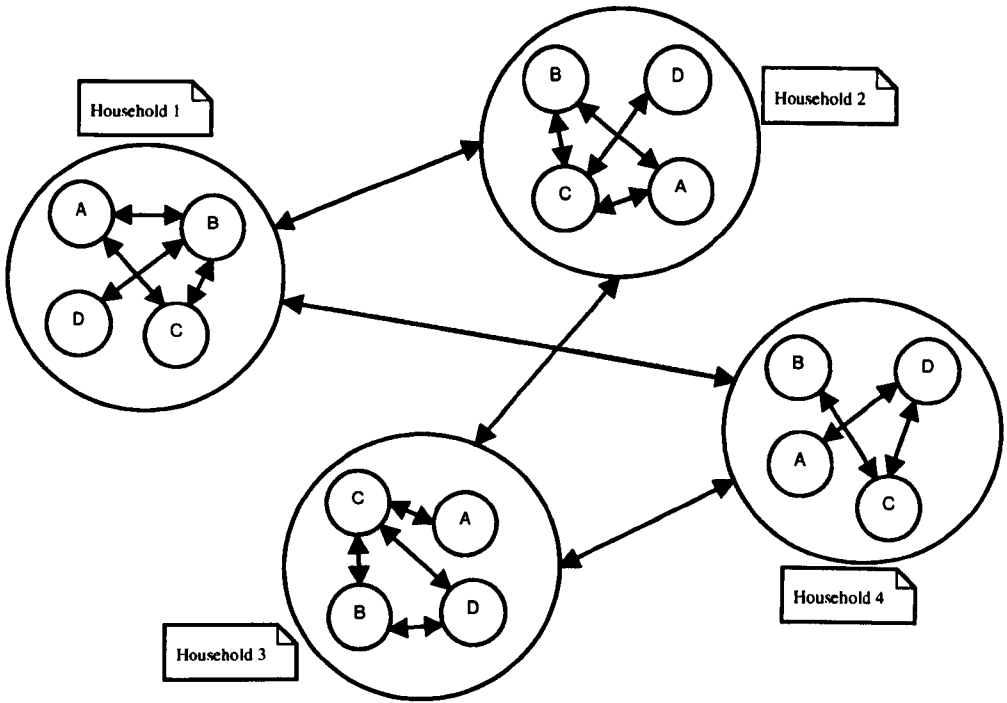


Figure 7.2: Coupling system used to represent the interaction between (blue arrow) and within (black arrow) households in the population. All households are assumed to contain the same number of individuals with each individual experiencing a higher rate of infection within the household compared to between households.

individuals in the population, we use the expression:

$$S = \frac{1}{N} \sum_{\forall a} a H_{a,b,c,d,e,f,g,h} \quad (7.1)$$

In the work presented in this chapter, we consider the situation where the number of individuals in a household, N , equals to 10. The stochastic nature of model is considered by modelling all the possible household configurations. The number of possible household configurations is dependent on the number of people in the household and the number of epidemiological classes under consideration. The

number of household configurations increases exponentially with both an increase in the number of epidemiological classes and household members. For example, assuming a modest $N = 10$, the number of household configurations increases from 66 with three classes to 19,448 with 8 epidemiological classes. This therefore requires that we keep both the number of people per household (N) and the epidemiological classes small although not too small as to compromise the natural history of the infection. The complete dynamics of the household model can be determined by considering the rates of movement between the epidemiological classes as represented by the ODE system shown in Eqn.(7.2). It is important to note that the impossible terms are excluded e.g. there can be no infection in a household with 10 infecteds or there can be no recovery in a household of completely susceptible individuals. For the sake of notational convenience, H will always represent $H_{a,b,c,d,e,f,g,h}$ unless otherwise stated. The integration of the resulting system of differential equations was solved numerically using MATLAB® [128]. The initial condition is such that 99% of the households are completely susceptible and the remaining 1% are in households with 2 infected individuals i.e.

1 in I_A and the other in I_B and the remaining in the susceptible class.

$$\begin{aligned}
\frac{dH}{dt} = & \frac{\beta_A}{N-1}(-abH + (a+1)(b-1)H_{a+1,b-1}) + \frac{\tau_A}{N-1}(-agH + (a+1)gH_{a+1,b-1}) \\
& + \frac{\beta_B}{N-1}(-acH + (a+1)(c-1)H_{a+1,c-1}) + \frac{\tau_B}{N-1}(-afH + (a+1)fH_{a+1,c-1}) \\
& + \gamma_A(-bH + (b+1)H_{b+1,d-1}) + \gamma_B(-cH + (c+1)H_{c+1,e-1}) \\
& + \frac{\sigma_{BA}\beta_B}{N-1}(-dcH + (d+1)cH_{d+1,f-1}) + \frac{\sigma_{BA}\tau_B}{N-1}(-dfH + (d+1)(f-1)H_{d+1,f-1}) \\
& + \frac{\sigma_{AB}\beta_A}{N-1}(-ebH + (e+1)bH_{e+1,g-1}) + \frac{\sigma_{AB}\tau_A}{N-1}(-egH + (e+1)(g-1)H_{e+1,g-1}) \\
& + \gamma_B(-fH + (f+1)H_{f+1,h-1}) + \gamma_A(-gH + (g+1)H_{g+1,h-1}) \\
& + \Omega_A(-hH + (h+1)H_{d-1,h+1}) + \Omega_B(-hH + (h+1)H_{e-1,h+1}) \\
& + \omega_A(-dH + (d+1)H_{a-1,d+1}) + \omega_B(-eH + (e+1)H_{a-1,e+1}) \\
& + \epsilon_B I_B + \alpha_B I_{BA}(-aH + (a+1)H_{a+1,c-1}) + \sigma_{BA}(\epsilon_B I_B + \alpha_B I_{BA})(-dH + (d+1)H_{d+1,f-1}) \\
& + \sigma_{AB}(\epsilon_A I_A + \alpha_A I_{AB})(-eH + (e+1)H_{e+1,g-1}) + \epsilon_A I_A + \alpha_A I_{AB}(-aH + (a+1)H_{a+1,b-1})
\end{aligned} \tag{7.2}$$

The cross immunity matrix is defined by $\begin{bmatrix} \sigma_{AA} & \sigma_{AB} \\ \sigma_{BA} & \sigma_{BB} \end{bmatrix}$ where σ_{AA} is the level of protection from re-infection with homologous A, σ_{AB} is the level of protection from re-infection with heterologous A, σ_{BA} is the level of protection from re-infection with heterologous B and σ_{BB} is the level of protection from re-infection with a homologous B. Table 7.1 gives the description and the values of the baseline parameters used in the simulation. The parameters presented are for demonstration purposes only in order to show the utility of such a model and should not be construed as necessarily accurate transmission parameters.

Table 7.1: Shows the model parameter explanation and their baseline values used in the multistrain model

Parameter	Description	Estimate
γ_A	Rate of recovery from A infection	1.2/week
γ_B	Rate of recovery from B infection	1.2/week
σ_{BA}	Cross protection from re-infection with B given a previous A	0.5
σ_{AB}	Cross protection from re-infection with A given a previous B	0.8
Ω_A	Rate of loss of protection from R_{AB} to R_A	2/week
Ω_B	Rate of loss of protection from R_{AB} to R_B	2/week
ω_A	Rate of loss of protection from R_A to S	0.05/week
ω_B	Rate of loss of protection from R_B to S	0.05/week
ϵ_A	External transmission parameter from I_A class	$0.6\beta_A$
α_A	External transmission parameter from I_{AB} class	$0.6\tau_A$
ϵ_B	External transmission parameter from I_B class	$0.6\beta_B$
α_B	External transmission parameter from I_{BA} class	$0.6\tau_B$
β_A	Household transmission parameter from I_A class	8/week
τ_A	Household transmission parameter from I_{AB} class	8/week
β_B	Household transmission parameter from I_B class	5/week
τ_B	Household transmission parameter from I_{BA} class	5/week

7.4 Results

In this section, we begin by presenting the equilibrium results at both the population and the household level. Figure 7.3 (A) shows the model run at equilibrium at the population level with the proportion of individuals with either RSV A or B shown on the y-axis and simulation time in years on the x-axis. We note that given the baseline parameters used in Table 7.1, the proportion of people with RSV A and B in the population is approximately equal at 0.22 of the entire population with majority of the infecteds in the secondary infected classes i.e. I_{AB} and I_{BA} as shown in Figure 7.3 (C). Figure 7.3 (B) and (D) shows the proportion of infected

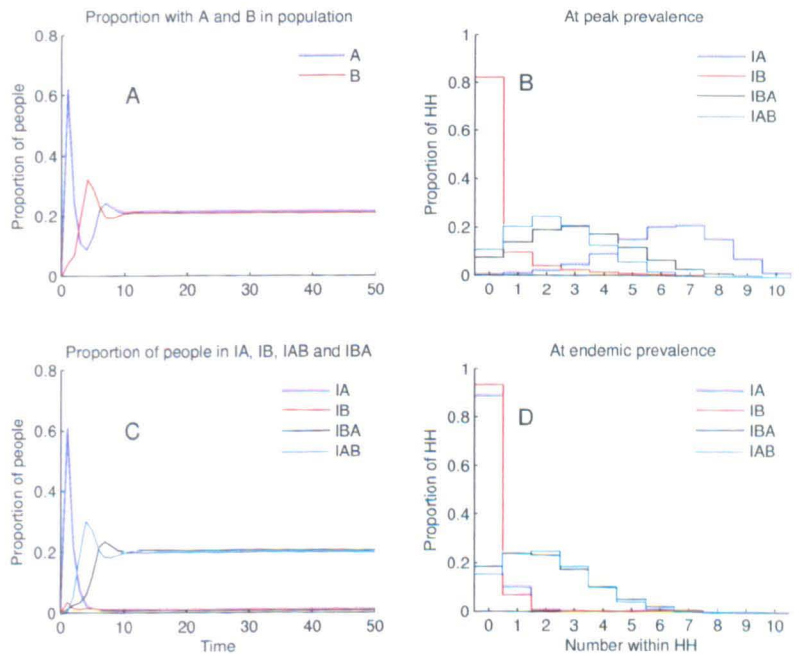


Figure 7.3: Epidemiological quantities as defined in the main text. A and C shows the transient and endemic equilibrium for the different infected classes at the population level while B and D shows the household dynamics at the peak and endemic prevalence respectively.

individuals in the households at both the peak and endemic prevalence. Most of the households, over 89%, at the endemic equilibrium as shown in Figure 7.3(D) have no infected individuals with neither primary RSV A nor B. The proportion of households, at endemic equilibrium, with over 2 individuals with primary infection (i.e. either RSV A or B) is very small at less than 0.002. However, at the same endemic equilibrium, the peak level of infection in households with secondary B infection is in households with 1 infected at 24% of the households while the peak in households with secondary A infection is in households with 2 infecteds at 25% of the distribution of households. The proportion of households decreases with an increase in the number of both secondary and primary infecteds.

Figure 7.4 shows the joint probability of observing co-infection in a household. Co-infection in a household is loosely defined as a the simultaneous infection of a household with both RSV A and B. The Figure 7.4 (A) and (B) shows the probability of co-infection with primary and secondary infections respectively. Let us suppose that x represents the number of individuals infected with RSV B and y represents the number of individuals infected with RSV A. Then Figure 7.4 (A) shows the the joint probability of observing x individuals with primary RSV B and y individuals with primary RSV A in a household i.e. $p(I_B = x \cup I_A = y)$ while subplot (B) shows $p(I_{BA} = x \cup I_{AB} = y)$. For primary infections, the highest joint probability is that of observing neither A nor B infections in a household at endemic equilibrium. However, for secondary infections, the joint probability is highest in households with 2 RSV A infections and 1 RSV B infection at endemic equilibrium. The white regions show the areas within which it is not possible to have a combination of infecteds since each household is assumed to have 10 individuals and therefore the joint probability is zero. This asymmetrical distribution

observed can be attributed to the interaction between the cross immunity term and the rate of infection.

At the endemic equilibrium as shown in Figure 7.3(C), the proportion of individuals in the population with secondary B infections (I_{BA}) is equal to that with secondary A infections. However, in the household distribution profile for infection with the two RSV groups, the peak level of infection with secondary B infections (I_{BA}) is in households with 1 infected while for secondary A infections (I_{AB}) is in households with 2 infected as can be seen in Figure 7.3 (D). This is an interesting outcome showing that, even though at the population level the infection appears indistinguishable, at the household level there is a clear difference in the peak of infection. The result suggests that, even though there is a strong interaction of the viruses within the household, it does not reflect at the population level.

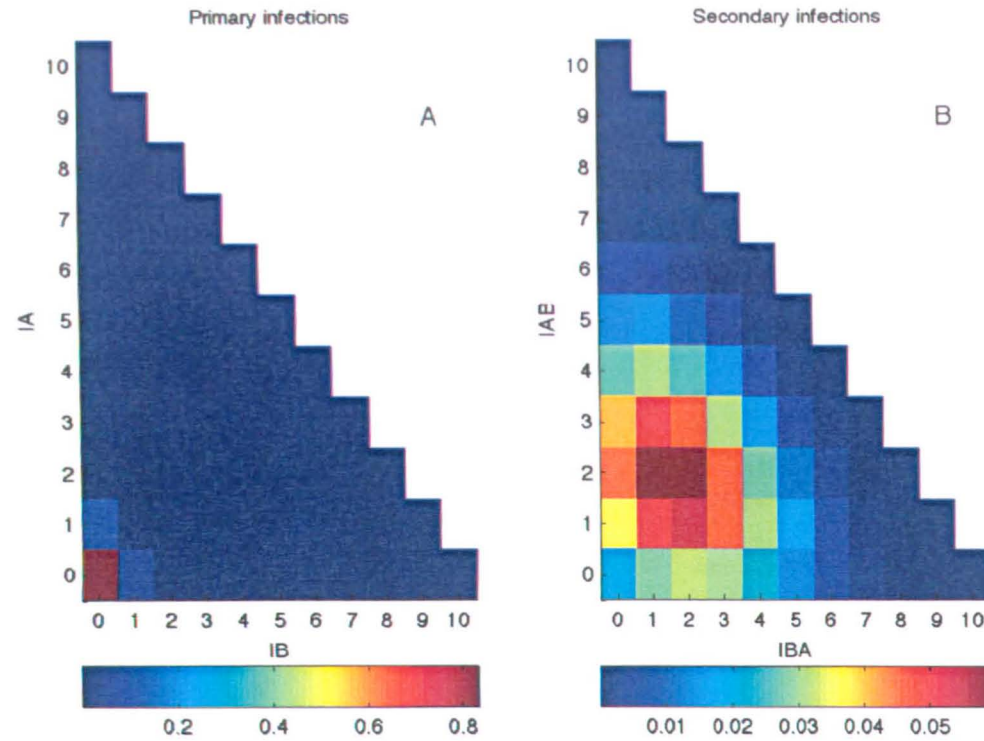


Figure 7.4: The joint probability of observing a household with x RSV B infections and y RSV A infections at equilibrium and for both primary and secondary infections.

7.5 Discussion

In this chapter, we have presented a mathematical model with two levels of transmission: at the household and at the population level. This model is obviously an over simplification of the complex dynamics of RSV transmission within the household and in the general population. The model presented is a very simple abstraction of the real infection dynamics and with the stochastic nature of transmission within the household accounted for by modelling all the possible household configurations. Our result shows that the general epidemic profile in the population may be different from the household epidemic profile as seen in Figure 7.3 (C) and (D). This allows for the household model to lend itself to the determination of more epidemiological parameters compared to simple mean field models.

The sort of modelling approach that we have taken has got some drawbacks. Firstly, even with a modest number of compartments, 8 epidemiological classes, and a fixed number of people in the households, 10 individuals, the system of ordinary differential equations ends up with 19,448 equations. Solving this system of ODE numerically takes a considerable amount of computer time on a desktop machine. However, there are computationally efficient methods that are being developed for estimating certain epidemiological quantities such as the threshold for invasion, endemic prevalence distribution and early growth rate [175, 44]. Secondly, we have only dealt with a population that is homogeneous i.e. where individuals and households are identical although we know that in reality households vary in size and in their composition. To accommodate this level of heterogeneity, we require some considerable level of model complexity which is the direction that we plan to extend this work. We plan to use this model to explore vac-

cination strategies targeted at households with both mono or bi-valent vaccines. Most importantly, we need to determine the model parameters especially both the household and the population wide transmission parameter and this can be done using a household infection data from a cohort that was monitored intensively with nasal samples taken every 3 to 4 days for RSV infection between January and June 2010 [140].

In general, we have highlighted some general concepts for epidemic modelling in populations that are structured into households and that can provide a basis for more parameters to be identified compared to simple ODE models. This kind of modelling can also be used to explore various vaccination strategies that are targeted towards households. However, more complex simulations models should be considered when determining policy.

Chapter 8

Final discussion

8.1 Introduction

The main findings arising from research presented in this thesis are summarized and discussed in this final chapter. The limitations of the work are highlighted and suggestions for future work have been presented.

8.2 Summary of the main findings

In the introductory chapter (Chapter 1), we set out the main objectives of the work presented in this thesis which were 1) to gain a better understanding of the transmission dynamics of RSV within a defined population, 2) estimate the social contact rate for the determination of the force of infection within the mathematical model framework and 3) to explore the impact of introducing routine vaccination in the population with an aim of identifying the optimal strategy in reducing the burden of disease in the infants.

In Chapter 4 we presented the results of the contact pattern within a Kenyan coastal population based on two data sets 1) data from the paper diary contact study and 2) data from the synthetic population based on household occupancy within the KHDSS and have considered the ability of the two datasets to accommodate the heterogeneous mixing patterns in the population. Besides the fact that these matrices have been constructed in different ways, they both share the same features i.e. strong assortative mixing by age among children with a decline in the strength in the older age groups. Another feature is the strong mixing between children and the individuals aged between 20 to 50 years, in the contact diary data, and 20 to 55 in the synthetic diary data. This most probably represents children having high rates of contact with their care takers at home, or if in school, with the teachers or instructors. The contact data as well reveals high mixing rates between infants and children aged 6 to 14 years. These 6 to 14 year olds are most likely to be primary school-going children. This sort of high inter-class mixing could be important in explaining the transmission of RSV to infants within the household. In fact, a previous household study identified that infants living in a household with at least one school-going child are at a higher risk of RSV infection compared to other infants in households with no school-going child [156]. This is further supported by a recent study that has demonstrated that older children, particularly school going, are the frequent introducers of RSV into households that lead to infant infection [140]. We also demonstrated that the pattern of mixing with the synthetic mixing data is mostly similar across different administrative locations within the KHDSS as seen in Figure 4.11. The only exception was in Junju and Kilifi Township where much of the mixing was recorded between adults approximately 20 years old. This may be due to households having fewer children

and the possibility of house sharing among adults. It also turns out that Kilifi Township and Junju have got the lowest mean number of household occupants at 6.7 and 7.9 respectively. If it is indeed true that households in these two regions are having fewer number of children that can be attributed to low fertility rates among adults living there, then it is possible Junju and Kilifi Township might be undergoing a demographic transition. It would be interesting to investigate what effect this kind of transition has on the transmission of infections. A direct comparison of the diary and the synthetic contact data by jointly linearly regressing their elements reveals that they are linearly correlated ($R^2 \approx 0.6$) implying a similarity in the mixing pattern observed. Similar results have been observed with a different formulation of the synthetic matrices from European countries [55, 60] which were then compared to POLYMOD contact diary data [136].

In Chapter 4 we have consistently assumed that the at-risk event of transmission of an infectious agent (in our case RSV) is a physical contact. The assertion is that if two individuals are close enough to touch, then they are probably close enough to transmit. This assertion is not necessarily true since it has been shown that RSV can also be transmitted through fomite [83] which can include sharing of objects. This would be true for people sharing public transport and they are exposed to touching communal surfaces or children at school sharing objects. The contacts are also assumed to be of equal intensity and hence the risk of transmission is homogeneous across all the contacts. However, contacts occurring within the household are likely to be more conducive to transmission since they last longer and are more intimate compared to contacts elsewhere [167]. In fact Edmunds et al [48] have suggested that one might imagine the relative risk of developing an infection from a contact in the different social settings obeys the following pattern:

home > work/social > background where background contacts are those that occur in shops or while travelling, with the background contacts acting as links to the more stable contacts occurring within the household or a social setting. However, we think that the data provides valuable information on the rates and pattern of mixing that may influence the spread of infectious diseases particularly RSV.

In Chapter 5, we then introduced a simple RSV mathematical model that was designed to enhance our understanding of the underlying RSV transmission process. The objectives were two-fold 1) to explore the stability of the model both at the disease free and the endemic equilibrium and 2) to explore conditions under which multiple endemic equilibria can exist i.e. the model exhibits a backward bifurcation curve. After analytically investigating the equilibrium properties of the model we found out that the disease free equilibrium is locally stable provided that $R_0 < 1$ and locally unstable when $R_0 = 1$. Given the difficulty in evaluating the model analytically beyond the invasion threshold i.e. at $R_0 > 1$, we ran numerical simulations that led to the observation that the model is locally stable when $R_0 > 1$. In the simplest form, backward bifurcation usually implies the existence of two subcritical endemic equilibria when $R_0 < 1$ and a unique supercritical endemic equilibrium when $R_0 > 1$. We have shown that multiple supercritical endemic equilibria exist as can be seen in Figure 5.5. An examination of the parameter set in Table 5.2 that resulted in the backward bifurcation shows that the parameter estimates are possibly realistic and this could have a bearing on the control of RSV. Classical epidemiological modelling implies that for the control of an infection, one needs to reduce R_0 to less than 1. This condition is however not necessarily satisfied with a model displaying backward bifurcation since there may exist endemic equilibria below $R_0 = 1$. This implies that for a disease

exhibiting a backward bifurcation curve, more effort is required to eliminate the infection from the population since you need to reduce the level of infection below that predicted by the invasion threshold. This analysis can be extended further by exploring whether the bifurcation curve can be linked in any way to the concept of probability of stochastic extinction. This would require the construction of a stochastic version of the model and to look at possible connections between the different bifurcation curves that might result and the stochastic phenomena observed.

We have further extended the model developed in Chapter 5 in order to include age heterogeneity within the transmission process in Chapter 6 with the aim of 1) developing a Realistic Age Structured (RAS) model reflecting the transmission dynamics of RSV within a defined population 2) evaluate the impact of routine RSV vaccination on the burden of disease in infants and 3) assess the variability of the vaccination programme outcome that is due to uncertainty in the model parameters using Latin Hypercube Sampling. We fitted the model to RSV hospitalization data from Kilifi using two mixing assumptions; diary and synthetic mixing assumptions. Fitting the model to temporal and age-specific RSV hospitalization data reveals that the model describes the data to a large extent since most of the model fit data points fall within the 95% CI of the hospitalization data.

It is clear that introduction of routine vaccination at any age between the first month of life and 5 years leads to a reduction in the number of hospitalizations without any adverse effects with the possibility of an increase in the average age at first infection for both mixing assumptions. At 7 months with 100% vaccination, the diary contact matrix model predicts that the average age at primary infection

will increase from 1.36 to approximately 6 years while that of the synthetic mixing matrix model increases from 1.7 to approximately 3 years with vaccination at 5 months with 100% vaccination coverage. If the development of disease is partly physiological, then older individuals who may have a higher residual for gaseous exchange may have a reduced risk of respiratory congestion such that an increase in the average age at first infection may be beneficial in reducing the amount of disease in the population and supported over early age classes by data from a birth cohort in Kilifi [155]. However, if disease does not decline with age in older age groups, or even possibly increasing, then the model may be overestimating the reduction in the burden of disease. However, the effect of this outcome remains unknown since, due to the ubiquitous nature of the virus, most of the first infections occur before the second year of life [68].

Both mixing assumptions have demonstrated that the effectiveness of the vaccine increases when delivery is delayed from birth as can be seen in Figures 6.7 and 6.15. For the model using the diary data, there is little decrease in the effectiveness when vaccination is delayed to between 5 and 12 months of age. In contrast, there is a sharp decline, beyond the fifth month of life, in the impact of vaccination on infant hospitalization using the model incorporating the synthetic mixing matrix. This observation seems to be the result of a trade off between the decay of protective level of RSV specific maternal antibody and the rate of infection. Although the synthetic contact matrix model predicts a longer duration of maternal antibodies, 3.6 months compared to 1.4 months for the diary-based contact matrix model, this does not translate to delayed optimal age at vaccination since as soon as children lose maternal protection, they are subjected to a higher rate of infection compared to the diary model, see Figure 6.10. On the other hand,

the diary matrix predicts a lower duration of maternal antibody protection hence the benefit of vaccinating early in life but little in the way of losing the benefit due to the lower rate of infection in the first year of life. However, this benefit is lost at the end of the first year of life when the rate of infection increases and this leads to a sharp decline in the benefit and this result is highly dependent upon the age-structure form of the force of infection. If these model results are robust, then a significant consequence is that the problem of vaccinating early in life can be circumvented. This may imply that much of the pressure in vaccine development could be lifted from early childhood vaccines and could accelerate the timetable to vaccine development, licensure and eventual introduction. However, given the variation in results of the two mixing matrices, the possibility of combining vaccination with measles vaccine delivery at 9 months remains uncertain.

Another interesting observation is that vaccination of children in the older age classes bring very little in the way of reducing the number of hospitalizations. This is in part because protecting the naive susceptibles will directly prevent disease since vaccinated individuals will not experience primary infection which is associated with increased risk of development of disease. Figures 6.4 and 6.6 show that the majority of the primary infections result from primary infections. However, we do not write-off school child vaccination since there may be benefit from vaccination across a wide age range at school rather than the delivery at a specific age gateway as currently specified in the RAS model. Furthermore, the model structure does not support household mixing which is required to create a strong mixing component between school going siblings and the infants. All the contacts in the model are treated with equal strength e.g. a contact occurring between an infant and an adult outside of the household contributes equally to the process of

infection with a similar contact within the household, despite the fact that household contacts are probably more stable temporally and more intimate [136, 166] and hence better placed to transfer the infection from one person to another.

The sensitivity and uncertainty analysis revealed that the duration of RSV specific maternal antibodies is important in explaining the variability in the number of age-specific hospitalizations and the optimal age at vaccination. An increase in the duration of the protective levels of maternal antibodies lead to an increase in the optimal age at which to vaccinate. This suggests that boosting the level of RSV specific maternal antibodies transfer from mother to child, increasing the duration of protection from infection of the child, would be another vaccination strategy that is potentially beneficial to explore and it is intended that such an analysis will be done. Longer protective duration would increase the average age at first infection therefore increasing the vaccination window to allow for vaccination in older children where vaccines have been shown to be both immunogenic and safe [114].

Even though the household model developed in Chapter 7 is an over simplification of reality, we have demonstrated that it can be used for the determination of epidemiologically relevant parameters that distinguish, at the household level, infection dynamics that are indistinguishable at the population level. Parameterizing this model requires the use of household infection data that is available to us through a recent cohort study that was carried out in the Kilifi KHDSS [140].

8.3 Limitations of the study and future research

During the course of the study, a number of limitations have been identified and that require further research. The limitations can be broadly categorized into two areas: 1) applications of the existing RSV model 2) improvements of the model with an aim of improving our biological and epidemiological understanding of RSV.

The structure of the contact matrix used to estimate the force of infection is of importance to the vaccination outcome. One of the main assumptions in using the diary contact data is that recorded physical contacts are the only at-risk events. However, there may exist other risk events that are important for the transmission of RSV e.g. being at close proximity with another person but not touching [16] and contacting items that are contaminated by the virus that are separated from the infecting individual [84]. The study was designed in such a way that it spanned six months in order to help account for any seasonal changes in contact patterns. We established, that potential seasonal triggers might be influenced by farming practices, fishing patterns and tourism [182]. During the farming season, most people will temporally leave their normal residential areas and move to work in the farms where they might have contacts with a different set of new individuals or even fewer contacts. However, even in our best attempt to account for seasonal changes, there is still a residual seasonal bias since we did not sample throughout an entire year. Therefore the extent to which the contact data should be interpreted as representative of the average throughout a year should be interpreted in view of this limitation. Another limitation of the diary study is that a majority of the participants ($\approx 60\%$) had their diaries filled in by a shadow. This is because those participants could neither adequately understand what was required to complete

the diary or had inadequate reading or writing skills. This applied either to children under 11 years and to the elderly. Shadowing has the disadvantage that it can lead to behaviour modification or failure to disclose all of the contacts. Recall bias is also another limitation that plagues diary studies and although every effort was made to reduce recall bias by encouraging participants to record their contacts as regularly as was practically possible and giving them watches that went off at a certain interval, it is likely that the data is still subject to this bias. The synthetic contact matrix on the other hand is based on data from household occupancy registers of the Kilifi HDSS. Assumptions about mixing outside of the household were made in order to estimate the population wide WAIFW matrix. In future, effort should be made to acquire data from other settings outside of the household e.g. from school attendance registers and the work place. The synthetic contact matrix approach can easily be extended to settings without social mixing data and can also be used to reconstruct contact patterns from the past by using previous census data, and this would be useful in identifying the influence of demographic transition on transmission of not only RSV but other respiratory pathogens requiring close contacts for their transmission.

The analysis in Chapter 6 has predominantly focused on implementing a delayed vaccination strategy in reducing the burden of disease in the infants. This analysis should be extended to examine the outcome of other vaccination strategies. This would be simple to achieve since the model structure has already been developed and the required data on social mixing patterns is available. Two forms of vaccination can be included 1) maternal vaccination and 2) post-natal vaccination. Maternal vaccination will have the benefit of boosting the level and duration of protection in the infant. Post natal vaccination can take the form of delayed

vaccination (as implemented in Chapter 6) or campaign vaccination over a wide age range. Another process of immunization that should be looked at in conjunction with the ones mentioned above is passive or prophylactic vaccination of high risk infants. High risk infants e.g. premature infants would be a relatively small number and hence unlikely to affect the transmission pattern. However, significant decrease in the overall burden of disease may be observable since vaccinating them directly protects them from primary infection and hence disease. The timing of the vaccination may be particularly important in this instance i.e. prophylactic vaccination would be best suited just before the beginning of the epidemic in order to boost the level of antibody protection at the time when most needed as opposed to giving it at birth.

Another limitation of the model is that the age specific fertility and mortality rates that we have used in our model are from a developing country setting. This challenges the extent to which our results can be generalized to a developed country setting. The model can be easily extended to make use of demographic data from the developed country setting and compare the vaccination outcome from the two settings. Finally, we have assumed that the vaccine elicits an immune response that is equivalent to a natural infection and that it protects against infection but with waning immunity. It would be of interest to evaluate the vaccine outcome when the vaccine elicits more or less protection compared to natural infection. Further, vaccination with a vaccine that protects against disease rather than infection would have a bearing in reducing the burden of disease but may have very little in the way of reducing the level of infection in the population and this would potentially reduce the strong indirect protection effects observed. We envisage to extend this model in the near future to accommodate more vaccination strategies and different

vaccine properties.

One of the dominant features about RSV is the regular seasonal epidemics. We have modelled these epidemics using a cosinusoidal function since the drivers of the seasonality are not known. There is therefore need to carry out studies trying to identify the seasonal triggers so as to facilitate their explicit inclusion in the model. However, it is worth noting that even with the cosinusoidal function, the model still captures the temporal epidemics. Another limitation of the work presented in Chapter 5 and 6 is that we have only implicitly accounted for antigenic differences between the two RSV strains i.e. group A and B. Work done by White et al. [206] has shown that the inclusion of the groups can in part explain the pattern of dominance of a single group during alternating epidemics. Additionally, the host immune response is likely to be important in estimating the outcome of vaccination especially given that a vaccine can have differential ability to generate an immunological response across the two groups or even sub-groups within a single group. However, this limitation would be mitigated if we consider a vaccine that targets the F-protein which is highly conserved between the two groups.

In the household modelling work, we have considered a homogeneous population i.e. where individuals and household are similar. We plan to extend this work by relaxing this assumptions and parameterizing the model using household cohort infection data from our setting. Using the model, we can then explore a vaccination strategy that targets households with a mono and/or a bi-valent vaccine.

8.4 Concluding remarks

In this thesis, we have taken a multidisciplinary approach in trying to understand the transmission dynamics of RSV and the impact of introducing vaccination. The mathematical model that we have *developed* has the capacity to capture the observed patterns in the epidemiology of RSV. There is currently no approved RSV vaccine and the goal of preventing RSV disease in the population therefore remains unmet. The success and choice of the immunization regime to adopt will be dependent on a number of factors which include epidemiological, logistical, economic and political goodwill.

Bibliography

- [1] ACEDO, L., DIEZ-DOMINGO, J., MORANO, J. A., AND VILLANUEVA, R. J. Mathematical modelling of respiratory syncytial virus (RSV): vaccination strategies and budget applications. *Epidemiol Infect* (2009), 1–8.
- [2] AGARWAL, R. P., AND O'REGAN, D. *An Introduction to Ordinary Differential Equations*. SpringerLink, 2008.
- [3] AGOTI, C. N., MWIHURI, A. G., SANDE, C., ONYANGO, C. O., MEDLEY, G. F., CANE, P. A., AND NOKES, J. D. Genetic relatedness of infecting and reinfecting respiratory syncytial virus strains identified in a birth cohort from rural Kenya. *Journal of Infectious Diseases, The* 206, 10 (2012), 1532–1541.
- [4] AKERLIND, B., NORRBY, E., ORVELL, C., AND MUFSON, M. A. Respiratory syncytial virus: heterogeneity of subgroup B strains. *The Journal of general virology* 69 (Pt 9) (Sept. 1988), 2145–54.
- [5] ANDERSON, L. J., DORMITZER, P. R., NOKES, D. J., RAPPUOLI, R., ROCA, A., AND GRAHAM, B. S. Strategic priorities for RSV vaccine development. *Vaccine In Press* (2013).
- [6] ANDERSON, L. J., HIERHOLZER, J. C., TSOU, C., HENDRY, R. M., FERNIE, B. F., STONE, Y., AND MCINTOSH, K. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *The Journal of infectious diseases* 151, 4 (Apr. 1985), 626–33.
- [7] ANDERSON, L. J., PARKER, R. A., STRIKAS, R. A., FARRAR, J. A., GANGAROSA, E. J., KEYSERLING, H. L., AND SIKES, R. K. Day-care center

- attendance and hospitalization for lower respiratory tract illness. *Pediatrics* 82, 3 (Sept. 1988), 300–8.
- [8] ANDERSON, R. M., AND MAY, R. M. Directly Transmitted Infectious Diseases: Control by Vaccination. *Science* 215, 4536 (1982), 1053–1060.
 - [9] ANDERSON, R. M., AND MAY, R. M. Age-related changes in the rate of disease transmission: implications for the design of vaccination programmes. *J Hyg (Lond)* 94, 3 (1985), 365–436.
 - [10] ANDERSON, R. M., AND MAY, R. M. *Infectious diseases of humans : dynamics and control*. Oxford University Press, Oxford; New York, 1991.
 - [11] AVENDAÑO, L. F., LARRAÑAGA, C., PALOMINO, M. A., GAGGERO, A., MONTALDO, G., SUÁREZ, M., AND DÍAZ, A. Community- and hospital-acquired respiratory syncytial virus infections in Chile. *The Pediatric infectious disease journal* 10, 8 (Aug. 1991), 564–8.
 - [12] BABAD, H. R., NOKES, D. J., GAY, N. J., MILLER, E., MORGAN-CAPNER, P., AND ANDERSON, R. M. Predicting the impact of measles vaccination in England and Wales: model validation and analysis of policy options. *Epidemiol Infect* 114, 2 (1995), 319–344.
 - [13] BALL, F. Stochastic and deterministic models for SIS epidemics among a population partitioned into households. *Mathematical biosciences* 156, 1-2 (Mar. 1999), 41–67.
 - [14] BECKER, N. G., AND DIETZ, K. The effect of household distribution on transmission and control of highly infectious diseases. *Math Biosci* 127, 2 (1995), 207–219.
 - [15] BERMAN, S., DUENAS, A., AND BEDOYA, A. Acute lower respiratory tract illnesses in Cali, Colombia: a two-year ambulatory study. *Pediatrics* 71, 2 (1983), 210–218.
 - [16] BEUTELS, P., SHKEDY, Z., AERTS, M., AND VAN DAMME, P. Social mixing patterns for transmission models of close contact infections: exploring

- self-evaluation and diary-based data collection through a web-based interface. *Epidemiology and infection* 134, 6 (Dec. 2006), 1158–66.
- [17] BLOWER, S., AND DOWLATABADI, H. Sensitivity and uncertainty analysis of complex models of disease transmission: an HIV model, as an example. *International Statistical Review/Revue Internationale de Statistique* 62, 2 (1994), 229–243.
 - [18] BLOWER, S. M., HARTEL, D., DOWLATABADI, H., ANDERSON, R. M., AND MAY, R. M. Drugs, sex and HIV: a mathematical model for New York City. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 331, 1260 (Feb. 1991), 171–87.
 - [19] BORRERO, I., FAJARDO, L., BEDOYA, A., ZEA, A., CARMONA, F., AND DE BORRERO, M. F. Acute respiratory tract infections among a birth cohort of children from Cali, Colombia, who were studied through 17 months of age. *Reviews of infectious diseases* 12 Suppl 8, December (1990), S950–6.
 - [20] BRANDENBURG, A. H., GROEN, J., VAN STEENSEL-MOLL, H. A., CLAAS, E. C., ROTHBARTH, P. H., NEIJENS, H. J., AND OSTERHAUS, A. D. Respiratory syncytial virus specific serum antibodies in infants under six months of age: limited serological response upon infection. *J Med Virol* 52, 1 (1997), 97–104.
 - [21] BRUHN, F. W., AND YEAGER, A. S. Respiratory syncytial virus in early *infancy*. Circulating antibody and the severity of infection. *Am J Dis Child* 131, 2 (1977), 145–148.
 - [22] BULKOW, L., SINGLETON, R., KARRON, R., AND HARRISON, L. Risk Factors for Severe Respiratory Syncytial Virus Infection Among Alaska Native Children. *Pediatrics* 109, 2 (2002), 210–6.
 - [23] CANE, P. A. Molecular epidemiology of respiratory syncytial virus. *Rev Med Virol* 11, 2 (2001), 103–16.

- [24] CANE, P. A., MATTHEWS, D. A., AND PRINGLE, C. R. Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *J Gen Virol* 72 (Pt 9) (1991), 2091–2096.
- [25] CANE, P. A., THOMAS, H. M., SIMPSON, A. F., EVANS, J. E., HART, C. A., AND PRINGLE, C. R. Analysis of the human serological immune response to a variable region of the attachment (G) protein of respiratory syncytial virus during primary infection. *Journal of medical virology* 48, 3 (Mar. 1996), 253–61.
- [26] CANE, P. A., WEBER, M., SANNEH, M., DACKOUR, R., PRINGLE, C. R., AND WHITTLE, H. Molecular epidemiology of respiratory syncytial virus in The Gambia. *Epidemiol Infect* 122, 1 (1999), 155–160.
- [27] CARBONELL-ESTRANY, X., AND QUERO, J. Hospitalization rates for respiratory syncytial virus infection in premature infants born during two consecutive seasons. *Pediatr Infect Dis J* 20, 9 (2001), 874–879.
- [28] CHAN, P., AND GOH, A. Respiratory Syncytial Virus Infection in Young Malaysian Children. *Singapore medical journal* 40, 5 (1999), 336–40.
- [29] CHAN, P., AND SUNG, R. Epidemiology of respiratory syncytial virus infection among paediatric patients in Hong Kong: seasonality and disease impact. *Epidemiology and ...* 123, 2 (Oct. 1999), 257–62.
- [30] CHANOCK, R., AND FINBERG, L. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). II. Epidemiologic aspects of infection in infants and young children. *American Journal of hygiene* 66, 3 (1957), 291–300.
- [31] CHANOCK, R., ROIZMAN, B., AND MYERS, R. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent (CCA). I. Isolation, properties and characterization. *American Journal of hygiene* 66, 3 (1957), 281–290.

- [32] CHEW, F. T., DORAISINGHAM, S., LING, A. E., KUMARASINGHE, G., AND LEE, B. W. Seasonal trends of viral respiratory tract infections in the tropics. *Epidemiol Infect* 121, 1 (1998), 121–128.
- [33] CHIN, J., MAGOFFIN, R. L., SHEARER, L. A., SCHIEBLE, J. H., AND LENNETTE, E. H. Field evaluation of a respiratory syncytial virus vaccine and a trivalent parainfluenza virus vaccine in a pediatric population. *Am J Epidemiol* 89, 4 (1969), 449–463.
- [34] CLINICALTRIAL.GOV. A Study to Evaluate the Safety, Tolerability, Immunogenicity and Vaccine-like Viral Shedding of MEDI-534, Against Respiratory Syncytial Virus (RSV) and Parainfluenza Virus Type 3 (PIV3), in Healthy 6 to < 24 Month-Old Children and in 2 Month Old Infants, 2008.
- [35] CLINICALTRIAL.GOV. A Randomized, Double-Blind, Placebo-Controlled Study to Evaluate Safety of MEDI-559 in Healthy 1 to <24 Month-Old Children, 2012.
- [36] COLLARINI, E. J., LEE, F. E.-H., FOORD, O., PARK, M., SPERINDE, G., WU, H., HARRIMAN, W. D., CARROLL, S. F., ELLSWORTH, S. L., ANDERSON, L. J., TRIPP, R. A., WALSH, E. E., KEYT, B. A., AND KAUVAR, L. M. Potent high-affinity antibodies for treatment and prophylaxis of respiratory syncytial virus derived from B cells of infected patients. *Journal of immunology* 183, 10 (Nov. 2009), 6338–45.
- [37] COX, M. J., AZEVEDO, R. S., CANE, P. A., MASSAD, E., AND MEDLEY, G. F. Seroepidemiological study of respiratory syncytial virus in Sao Paulo state, Brazil. *J Med Virol* 55, 3 (1998), 234–239.
- [38] CROWE JR., J. E. Influence of maternal antibodies on neonatal immunization against respiratory viruses. *Clin Infect Dis* 33, 10 (2001), 1720–7.
- [39] CULLEY, F. J., POLLOTT, J., AND OPENSHAW, P. J. Age at First Viral Infection Determines the Pattern of T Cell-mediated Disease during Reinfection in Adulthood. *Journal of Experimental Medicine* 196, 10 (Nov. 2002), 1381–1386.

- [40] DELVALLE, S., HYMAN, J., HETHCOTE, H., AND EUBANK, S. Mixing patterns between age groups in social networks. *Social Networks* 29, 4 (Oct. 2007), 539–554.
- [41] DEVINCENZO, J. P., WILKINSON, T., VAISHNAW, A., CEHELKY, J., MEYERS, R., NOCHUR, S., HARRISON, L., MEEKING, P., MANN, A., MOANE, E., OXFORD, J., PAREEK, R., MOORE, R., WALSH, E., STUDHOLME, R., DORSETT, P., ALVAREZ, R., AND LAMBKIN-WILLIAMS, R. Viral load drives disease in humans experimentally infected with respiratory syncytial virus. *American Journal of Respiratory and Critical Care Medicine* 182, 10 (Nov. 2010), 1305–14.
- [42] DIEKMANN, O., HEESTERBEEK, J. A. P., AND METZ, J. A. J. On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *Journal of Mathematical Biology* 28, 4 (1990), 365–382.
- [43] DJELANTIK, I. G., GESSNER, B. D., SOEWIGNJO, S., STEINHOFF, M., SUTANTO, A., WIDJAYA, A., LINEHAN, M., MONIAGA, V., AND INGERANI. Incidence and clinical features of hospitalization because of respiratory syncytial virus lower respiratory illness among children less than two years of age in a rural Asian setting. *The Pediatric Infectious Disease Journal* 22, 2 (Feb. 2003), 150–7.
- [44] DODD, P. J., AND FERGUSON, N. M. Approximate disease dynamics in household-structured populations. *Journal of the Royal Society, Interface / the Royal Society* 4, 17 (Dec. 2007), 1103–6.
- [45] DORIGATTI, I., AND PUGLIESE, A. Analysis of a vaccine model with cross-immunity: When can two competing infectious strains coexist? *Mathematical Biosciences* 234, 1 (Nov. 2011), 33–46.
- [46] EARN, D. J. D., ROHANI, P., BOLKER, B. M., AND GRENFELL, B. T. A Simple Model for Complex Dynamical Transitions in Epidemics. *Science* 287, 5453 (Jan. 2000), 667–670.

- [47] EDMUNDS, W. J., MEDLEY, G. F., AND NOKES, D. J. The transmission dynamics and control of hepatitis B virus in The Gambia. *Stat Med* 15, 20 (1996), 2215–2233.
- [48] EDMUNDS, W. J., O'CALLAGHAN, C. J., AND NOKES, D. J. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proceedings. Biological sciences / The Royal Society* 264, 1384 (July 1997), 949–57.
- [49] EICK, A., KARRON, R., SHAW, J., THUMAR, B., REID, R., SANTOSHAM, M., AND O'BRIEN, K. L. The role of neutralizing antibodies in protection of American Indian infants against respiratory syncytial virus disease. *The Pediatric infectious disease journal* 27, 3 (Mar. 2008), 207–12.
- [50] ENGLUND, J., PAUL GLEZEN, W., AND PIEDRA, P. A. Maternal immunization against viral disease. *Vaccine* 16, 14-15 (1998), 1456–1463.
- [51] ENGLUND, J. A. Passive protection against respiratory syncytial virus disease in infants: the role of maternal antibody. *The Pediatric infectious disease journal* 13, 5 (May 1994), 449–53.
- [52] FALSEY, A., HENNESSEY, P., FORMICA, M., CHRISTOPHER, C., AND WALSH, E. Respiratory Syncytial Virus Infection in Elderly and High-Risk Adults. *New England Journal of Medicine* 325, 17 (2005), 1749–1759.
- [53] FALSEY, A. R., CUNNINGHAM, C. K., BARKER, W. H., KOUIDES, R. W., YUEN, J. B., MENEGUS, M., WEINER, L. B., BONVILLE, C. A., AND BETTS, R. F. Respiratory syncytial virus and influenza A infections in the hospitalized elderly. *The Journal of infectious diseases* 172, 2 (Aug. 1995), 389–94.
- [54] FARRINGTON, C., KANAAN, M., AND GAY, N. J. Estimation of the basic reproduction number for infectious diseases from age-stratified serological survey data. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 50, 3 (2001), 25–292.

- [55] FAVA, E. D. *Statistical methods for modeling of drug-related and close-contact infections*. PhD thesis, Hasselt University, 2012.
- [56] FERGUSON, N. M., ANDERSON, R. M., AND GARNETT, G. P. Mass vaccination to control chickenpox: the influence of zoster. *Proceedings of the National Academy of Sciences* 93, 14 (1996), 7231–7235.
- [57] FINE, P. E., AND CLARKSON, J. A. Measles in England and Wales–II: The impact of the measles vaccination programme on the distribution of immunity in the population. *Int J Epidemiol* 11, 1 (1982), 15–25.
- [58] FOX, J. P., ELVEBACK, L., SCOTT, W., GATEWOOD, L., AND ACKERMAN, E. Herd immunity: basic concept and relevance to public health immunization practices. 1971. *Am J Epidemiol* 141, 3 (1995), 186–187.
- [59] FRANK, A. L., TABER, L. H., WELLS, C. R., WELLS, J. M., GLEZEN, W. P., AND A, P. Patterns of shedding of myxoviruses and paramyxoviruses in children. *Journal of Infectious Diseases, The* 144, 5 (1981), 433–441.
- [60] FUMANELLI, L., AJELLI, M., MANFREDI, P., VESPIGNANI, A., AND MERLER, S. Inferring the Structure of Social Contacts from Demographic Data in the Analysis of Infectious Diseases Spread. *PLoS computational biology* 8, 9 (Sept. 2012), e1002673.
- [61] GALEA, S., AND TRACY, M. Participation Rates in Epidemiologic Studies. *Annals of Epidemiology* 17, 9 (Sept. 2007), 643–653.
- [62] GHOSHAL, G., SANDER, L. M., AND SOKOLOV, I. M. SIS epidemics with household structure: the self-consistent field method. *Mathematical biosciences* 190, 1 (July 2004), 71–85.
- [63] GILLESPIE, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry* 81, 25 (1977), 2340–2361.
- [64] GLENN, G. M., SMITH, G., FRIES, L., RAGHUNANDAN, R., LU, H., ZHOU, B., THOMAS, D. N., HICKMAN, S. P., KPAMEGAN, E., BODDAPATI, S., AND PIEDRA, P. A. Safety and immunogenicity of a Sf9 insect

cell-derived respiratory syncytial virus fusion protein nanoparticle vaccine. *Vaccine* 31, 3 (Jan. 2013), 524–32.

- [65] GLEZEN, W. P., AND ALPERS, M. Maternal immunization. *Clin Infect Dis* 28, 2 (1999), 219–224.
- [66] GLEZEN, W. P., GAGLANI, M. J., KOZINETZ, C. A., AND PIEDRA, P. A. Direct and Indirect Effectiveness of Influenza Vaccination Delivered to Children at School Preceding an Epidemic Caused by 3 New Influenza Virus Variants. *The Journal of infectious diseases* (Oct. 2010).
- [67] GLEZEN, W. P., PAREDES, A., ALLISON, J. E., TABER, L. H., AND FRANK, A. L. Risk of respiratory syncytial virus infection for infants from low-income families in relationship to age, sex, ethnic group, and maternal antibody level. *The Journal of pediatrics* 98, 5 (May 1981), 708–15.
- [68] GLEZEN, W. P., TABER, L. H., FRANK, A. L., AND KASEL, J. A. Risk of primary infection and reinfection with respiratory syncytial virus. *American Journal of Diseases of Children* 140 (1986), 543–546.
- [69] GOMES, M. G., FRANCO, A. O., GOMES, M. C., AND MEDLEY, G. F. The reinfection threshold promotes variability in tuberculosis epidemiology and vaccine efficacy. *Proc Biol Sci* 271, 1539 (2004), 617–623.
- [70] GONZALEZ, P., BUENO, S., CARRENO, L., RIEDEL, C., AND KALERGIS, A. Respiratory syncytial virus infection and immunity. *Reviews in Medical Virology* (2012).
- [71] GRASSLY, N. C., AND FRASER, C. Seasonal infectious disease epidemiology. *Proceedings. Biological sciences / The Royal Society* 273, 1600 (Oct. 2006), 2541–50.
- [72] GRASSLY, N. C., AND FRASER, C. Mathematical models of infectious disease transmission. *Nat Rev Micro* 6, 6 (2008), 477–487.
- [73] GREENHALGH, D., DIEKMANN, O., AND DE JONG, M. C. M. Subcritical endemic steady states in mathematical models for animal infections with incomplete immunity. *Mathematical Biosciences* 165, 1 (2000), 1–25.

- [74] GREENHALGH, D., AND GRIFFITHS, M. Backward bifurcation, equilibrium and stability phenomena in a three-stage extended BRSV epidemic model. *J Math Biol* 59, 1 (2009), 1–36.
- [75] GREENHALGH, D., AND GRIFFITHS, M. Dynamic phenomena arising from an extended Core Group model. *Mathematical Biosciences* 221, 2 (2009), 136–149.
- [76] GROOTHUIS, J., SIMOES, E., MYRON, J., HALL, C., LONG, C. E., RODRIGUES, W., ARROBIO, J., MEISSNER, H., FULTON, D., WELLIVER, R. C., TRISTRAM, D., SIBER, G., PRINCE, G., RADEN, M., AND HEMMING, V. Prophylactic administration of respiratory syncytial virus immune globulin in high-risk infants and young children. *New England Journal of Medicine* 329, 21 (1993), 1524–30.
- [77] GÜRKAN, F., KIRAL, A., DAĞLI, E., AND KARAKOÇ, F. The effect of passive smoking on the development of respiratory syncytial virus bronchiolitis. *European journal of epidemiology* 16, 5 (May 2000), 465–8.
- [78] HACIMUSTAFAOĞLU, M., CELEBI, S., AYNACI, E., SINIRTAS, M., KOKSAL, N., KUCUKERDOĞAN, A., ERCAN, I., GORAL, G., AND ILDIRIM, I. The progression of maternal RSV antibodies in the offspring. *Archives of Disease in Childhood* 89, 1 (2004), 52–53.
- [79] HADELER, K. P., AND CASTILLO-CHAVEZ, C. A core group model for disease transmission. *Mathematical Biosciences* 128, 1-2 (1995), 41–55.
- [80] HALL, C. B. The shedding and spreading of respiratory syncytial virus. *Pediatr Res* 11, 3 Pt 2 (1977), 236–239.
- [81] HALL, C. B., AND DOUGLAS, R. G. Clinically useful method for the isolation of respiratory syncytial virus. *The Journal of infectious diseases* 131, 1 (Jan. 1975), 1–5.
- [82] HALL, C. B., DOUGLAS, R. G., AND GEIMAN, J. M. Quantitative shedding patterns of respiratory syncytial virus in infants. *The Journal of infectious diseases* 132, 2 (Aug. 1975), 151–6.

- [83] HALL, C. B., DOUGLAS, R. G., AND GEIMAN, J. M. Possible transmission by fomites of respiratory syncytial virus. *The Journal of infectious diseases* 141, 1 (Jan. 1980), 98–102.
- [84] HALL, C. B., DOUGLAS, R. G., SCHNABEL, K. C., AND GEIMAN, J. M. Infectivity of respiratory syncytial virus by various routes of inoculation. *Infection and immunity* 33, 3 (Sept. 1981), 779–83.
- [85] HALL, C. B., DOUGLAS, R. G. J., AND GEIMAN, J. M. Respiratory syncytial virus infections in infants: quantitation and duration of shedding. *Journal of Pediatrics* 89 (1976), 11–15.
- [86] HALL, C. B., AND DOUGLAS JR., R. G. Modes of transmission of respiratory syncytial virus. *J Pediatr* 99, 1 (1981), 100–103.
- [87] HALL, C. B., GEIMAN, J. M., BIGGAR, R., KOTOK, D. I., HOGAN, P. M., AND DOUGLAS JR., G. R. Respiratory syncytial virus infections within families. *N Engl J Med* 294, 8 (1976), 414–9.
- [88] HALL, C. B., LONG, C. E., AND SCHNABEL, K. C. Respiratory Syncytial Virus Infections in Previously Healthy Working Adults. *Clinical Infectious Diseases* 33, 6 (2001), 792–796.
- [89] HALL, C. B., WALSH, E. E., LONG, C. E., AND SCHNABEL, K. C. Immunity to and frequency of reinfection with respiratory syncytial virus. *The Journal of infectious diseases* 163, 4 (Apr. 1991), 693–8.
- [90] HAN, L. L., ALEXANDER, J. P., AND ANDERSON, L. J. Respiratory syncytial virus pneumonia among the elderly: an assessment of disease burden. *The Journal of infectious diseases* 179, 1 (Jan. 1999), 25–30.
- [91] HANSON, L. A. Breastfeeding provides passive and likely long-lasting active immunity. *Annals of allergy, asthma & immunology : official publication of the American College of Allergy, Asthma, & Immunology* 81, 6 (Dec. 1998), 523–33; quiz 533–4, 537.

- [92] HANSON, L. A., KOROTKOVA, M., LUNDIN, S., HÅ VERSEN, L., SILFVERDAL, S.-A., MATTSBY-BALTZER, I., STRANDVIK, B., AND TELEMO, E. The transfer of immunity from mother to child. *Annals of the New York Academy of Sciences* 987 (Apr. 2003), 199–206.
- [93] HENDERSON, F. W., COLLIER, A. M., CLYDE JR., W. A., AND DENNY, F. W. Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children. *N Engl J Med* 300, 10 (1979), 530–534.
- [94] HENDLEY, J., WENZEL, R., AND GWALTNEY, J. J. Transmission of rhinovirus colds by self-inoculation. *New England Journal of Medicine* 288, 28 (1973), 1361–4.
- [95] HENS, N., AYELE, G. M., GOEYVAERTS, N., AERTS, M., MOSSONG, J., EDMUNDS, J. W., AND BEUTELS, P. Estimating the impact of school closure on social mixing behaviour and the transmission of close contact infections in eight European countries. *BMC infectious diseases* 9 (Jan. 2009), 187.
- [96] HETHCOTE, H. W. An age-structured model for pertussis transmission. *Mathematical Biosciences* 145, 2 (1997), 89–136.
- [97] HETHCOTE, H. W. Simulations of pertussis epidemiology in the United States: effects of adult booster vaccinations. *Mathematical Biosciences* 158, 1 (1999), 47–73.
- [98] HIJAZI, Z., PACSA, A., EISA, S., EL SHAZLI, A., ABD EL SALAM, R., AND EL-GHARBAWY, F. Respiratory syncytial virus infections in children in a desert country. *The Pediatric infectious disease journal* 14, 4 (Apr. 1995), 322–4.
- [99] HOLBERG, C. J., WRIGHT, A. L., MARTINEZ, F. D., RAY, C. G., TAUSIG, L. M., AND LEBOWITZ, M. D. Risk factors for respiratory syncytial virus-associated lower respiratory illnesses in the first year of life. *American journal of epidemiology* 133, 11 (June 1991), 1135–51.

- [100] HORBY, P., PHAM, Q. T., HENS, N., NGUYEN, T. T. Y., LE, Q. M., DANG, D. T., NGUYEN, M. L., NGUYEN, T. H., ALEXANDER, N., EDMUNDS, W. J., TRAN, N. D., FOX, A., AND NGUYEN, T. H. Social contact patterns in Vietnam and implications for the control of infectious diseases. *PloS one* 6, 2 (Jan. 2011), e16965.
- [101] HOUSE, T., AND KEELING, M. J. Deterministic epidemic models with explicit household structure. *Math Biosci* 213, 1 (2008), 29–39.
- [102] HOUSE, T., AND KEELING, M. J. Household structure and infectious disease transmission. *Epidemiol Infect* 137, 5 (2009), 654–661.
- [103] HURWITZ, E. S., GUNN, W. J., PINSKY, P. F., AND SCHONBERGER, L. B. Risk of respiratory illness associated with daycare attendance a nationwide study. *Pediatrics* 87, 1 (Jan. 1991), 62–9.
- [104] IMAN, R. L., AND HELTON, J. C. An Investigation of Uncertainty and Sensitivity Analysis Techniques for Computer Models. *Risk Analysis* 8, 1 (Mar. 1988), 71–90.
- [105] IRVINE, D. J. Drug delivery: One nanoparticle, one kill. *Nature materials* 10, 5 (May 2011), 342–3.
- [106] JAMJOOM, G. A., AL-SEMRANI, A. M., BOARD, A., AL-FRAYH, A. R., ARTZ, F., AND AL-MOBAIREEK, K. F. Respiratory syncytial virus infection in young children hospitalized with respiratory illness in Riyadh. *Journal of tropical pediatrics* 39, 6 (Dec. 1993), 346–9.
- [107] JOHNSON, P., AND COLLINS, P. The Fusion Glycoproteins of Human Respiratory Syncytial Virus of Subgroups A and B: Sequence Conservation Provides a Structural Basis for Antigenic Relatedness. *The Journal of general virology* 2, 69 (1988), 2623–2628.
- [108] JOHNSON, P. R., OLMSTED, R. A., PRINCE, G. A., MURPHY, B. R., ALLING, D. W., WALSH, E. E., AND COLLINS, P. L. Antigenic relatedness between glycoproteins of human respiratory syncytial virus subgroups A and

- B: evaluation of the contributions of F and G glycoproteins to immunity. *Journal of virology* 61, 10 (Oct. 1987), 3163–6.
- [109] JOHNSON, P. R., SPRIGGS, M. K., OLMSTED, R. A., AND COLLINS, P. L. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proceedings of the National Academy of Sciences of the United States of America* 84, 16 (Aug. 1987), 5625–9.
 - [110] JOHNSTONE-ROBERTSON, S. P., MARK, D., MORROW, C., MIDDELKOOP, K., CHISWELL, M., AQUINO, L. D. H., BEKKER, L.-G., AND WOOD, R. Social Mixing Patterns Within a South African Township Community: Implications for Respiratory Disease Transmission and Control. *American journal of epidemiology* 174, 11 (Nov. 2011), 1246–1255.
 - [111] JOOSTING, A. C., HARWIN, R. M., ORCHARD, M., MARTIN, E., AND GEAR, J. H. Respiratory viruses in hospital patients on the Witwatersrand. A 7-year study. *South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde* 55, 11 (Mar. 1979), 403–8.
 - [112] KAPIKIAN, A. Z., BELL, J. A., MASTROTA, F. M., JOHNSON, K. M., HUEBNER, R. J., AND CHANOCK, R. M. An outbreak of febrile illness and pneumonia associated with respiratory syncytial virus infection. *American Journal of Epidemiology* 74, 3 (Nov. 1961), 234–48.
 - [113] KARRON, R. A., SINGLETON, R. J., BULKOW, L., PARKINSON, A., KRUSE, D., DESMET, I., INDORF, C., PETERSEN, K. M., LEOMBRUNO, D., HURLBURT, D., SANTOSHAM, M., AND HARRISON, L. H. Severe Respiratory Syncytial Virus Disease in Alaska Native Children. *Journal of Infectious Diseases, The* 180, 1 (1999), 41–49.
 - [114] KARRON, R. A., WRIGHT, P. F., BELSHE, R. B., THUMAR, B., CASEY, R., NEWMAN, F., POLACK, F. P., RANDOLPH, V. B., DEATLY, A., HACKELL, J., GRUBER, W., MURPHY, B. R., AND COLLINS, P. L. Identification of a recombinant live attenuated respiratory syncytial virus vaccine

- candidate that is highly attenuated in infants. *J Infect Dis* 191, 7 (2005), 1093–1104.
- [115] KEELING, M. The implications of network structure for epidemic dynamics. *Theoretical Population Biology* 67, 1 (2005), 1–8.
 - [116] KEELING, M. J., AND ROHANI, P. *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press, 2007.
 - [117] KRETZSCHMAR, M., MANGEN, M.-J., VAN DE LAAR, M., AND DE WIT, A. Model based analysis of hepatitis B vaccination strategies in the Netherlands. *Vaccine* 27, 8 (Feb. 2009), 1254–60.
 - [118] LAMPRECHT, C. L., KRAUSE, H. E., AND MUFSON, M. A. Role of Maternal Antibody in Pneumonia and Bronchiolitis Due to Respiratory Syncytial Virus. *Journal of Infectious Diseases* 134, 3 (Sept. 1976), 211–217.
 - [119] LESSLER, J., REICH, N. G., BROOKMEYER, R., PERL, T. M., NELSON, K. E., AND CUMMINGS, D. A. T. Incubation periods of acute respiratory viral infections: a systematic review. *The Lancet infectious diseases* 9, 5 (May 2009), 291–300.
 - [120] LIESE, J. G., GRILL, E., FISCHER, B., ROECKL-WIEDMANN, I., CARR, D., AND BELOHRADSKY, B. H. Incidence and risk factors of respiratory syncytial virus-related hospitalizations in premature infants in Germany. *European journal of pediatrics* 162, 4 (Apr. 2003), 230–6.
 - [121] LIPSITCH, M., COHEN, T., COOPER, B., ROBINS, J. M., MA, S., JAMES, L., GOPALAKRISHNA, G., CHEW, S. K., TAN, C. C., SAMORE, M. H., FISMAN, D., AND MURRAY, M. Transmission Dynamics and Control of Severe Acute Respiratory Syndrome. *Science* 300, 5627 (2003), 1966–1970.
 - [122] LONGINI JR., I. M., KOOPMAN, J. S., MONTO, A. S., AND FOX, J. P. Estimating household and community transmission parameters for influenza. *Am J Epidemiol* 115, 5 (1982), 736–751.

- [123] LÓPEZ-ALARCÓN, M., VILLALPANDO, S., AND FAJARDO, A. Breast-feeding lowers the frequency and duration of acute respiratory infection and diarrhea in infants under six months of age. *The Journal of Nutrition* 127, 3 (1997), 436–443.
- [124] LOPMAN, B. A., CURNS, A. T., YEN, C., AND PARASHAR, U. D. Infant rotavirus vaccination may provide indirect protection to older children and adults in the United States. *The Journal of infectious diseases* 204, 7 (Oct. 2011), 980–6.
- [125] LUZ, P. M., CODECO, C. T., WERNECK, G. L., AND STRUCHINER, C. J. A modelling analysis of pertussis transmission and vaccination in Rio de Janeiro, Brazil. *Epidemiol Infect* 134, 4 (2006), 850–862.
- [126] MARINO, S., HOGUE, I., RAY, C., AND KIRSCHNER, D. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of theoretical biology* 254, 1 (2008), 178–196.
- [127] MARTIN, A., GARDNER, P., AND MCQUILLIN, J. No Epidemiology of respiratory viral infection among paediatric inpatients over a six year period in north-east England. *Lancet* 11, 2 (1978), 1038–8.
- [128] MATHWORKS. MATLAB, 2008.
- [129] MCKAY, M., BECKMAN, R., AND CONOVER, W. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 2 (1979), 239–245.
- [130] MCLEAN, A. R., AND BLUMBERG, B. S. Modelling the Impact of Mass Vaccination against Hepatitis B. I. Model Formulation and Parameter Estimation. *Proceedings: Biological Sciences* 256, 1345 (1994), 7–15.
- [131] MEDLEY, G. F., LINDOP, N. A., EDMUNDS, W. J., AND NOKES, D. J. Hepatitis-B virus endemicity: heterogeneity, catastrophic dynamics and control. *Nat Med* 7, 5 (2001), 619–624.

- [132] MEIJBOOM, M. J., ROZENBAUM, M. H., BENEDICTUS, A., LUYTJES, W., KNEYBER, M. C. J., WILSCHUT, J. C., HAK, E., AND POSTMA, M. J. Cost-effectiveness of potential infant vaccination against respiratory syncytial virus infection in The Netherlands. *Vaccine* 30, 31 (June 2012), 4691–700.
- [133] MIKOLAJCZYK, R. T., AKMATOV, M. K., RASTIN, S., AND KRETZSCHMAR, M. Social contacts of school children and the transmission of respiratory-spread pathogens. *Epidemiology and Infection* 136, 06 (2008), 813–822.
- [134] MONTA, A. S., AND LIM, S. K. THE TECUMSEH STUDY OF RESPIRATORY ILLNESS: III. INCIDENCE AND PERIODICITY OF RESPIRATORY SYNCYTIAL VIRUS AND MYCOPLASMA PNEUMONIAE INFECTIONS. *American Journal of Epidemiology* 94, 3 (1971), 290–301.
- [135] MORRIS, J., BLOUNT, R., AND SAVAGE, R. Recovery of Cytopathogenic Agent from Chimpanzees with coryza. *Proceedings of the Society for Experimental Biology and Medicine* 92, 3 (1956), 544–549.
- [136] MOSSONG, J., HENS, N., JIT, M., BEUTELS, P., AURANEN, K., MIKOLAJCZYK, R., MASSARI, M., SALMASO, S., TOMBA, G. S., WALLINGA, J., HEIJNE, J., SADKOWSKA-TODYS, M., ROSINSKA, M., AND EDMUNDS, W. J. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 5, 3 (2008), e74.
- [137] MUFSON, M. A., BELSHE, R. B., ORVELL, C., AND NORRBY, E. Subgroup characteristics of respiratory syncytial virus strains recovered from children with two consecutive infections. *Journal of clinical microbiology* 25, 8 (Aug. 1987), 1535–9.
- [138] MUFSON, M. A., ORVELL, C., RAFNAR, B., AND NORRBY, E. Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol* 66 (1985), 2111–2124.

- [139] MUNOZ, F. M., PIEDRA, P. A., AND GLEZEN, W. P. Safety and immunogenicity of respiratory syncytial virus purified fusion protein-2 vaccine in pregnant women. *Vaccine* 21, 24 (2003), 3465–3467.
- [140] MUNYWOKI, P., KOECH, D., AGOTI, C. N., LEWA, C., CANE, P., MEDLEY, G. F., AND NOKES, D. J. The source of respiratory syncytial virus infection in infants: a household cohort study in rural Kenya. *In Press*.
- [141] MUNYWOKI, P. K., HAMID, F., MUTUNGA, M., WELCH, S., CANE, P., AND NOKES, D. J. Improved detection of respiratory viruses in pediatric outpatients with acute respiratory illness by real-time PCR using nasopharyngeal flocked swabs. *Journal of clinical microbiology* 49, 9 (Sept. 2011), 3365–7.
- [142] MUNYWOKI, P. K., MEDLEY, G. F., AND NOKES, D. J. RSV shedding patterns. *To be submitted* (2013).
- [143] MURRAY, C. S., PIPIS, S. D., MCARDLE, E. C., LOWE, L. A., CUSTOVIC, A., AND WOODCOCK, A. Lung function at one month of age as a risk factor for infant respiratory symptoms in a high risk population. *Thorax* 57, 5 (May 2002), 388–92.
- [144] MYUNG, I. J. Tutorial on maximum likelihood estimation. *Journal of Mathematical Psychology* 47, 1 (2003), 90–100.
- [145] NAIR, H., NOKES, D. J., GESSNER, B. D., DHERANI, M., MADHI, S. A., SINGLETON, R. J., O'BRIEN, K. L., ROCA, A., WRIGHT, P. F., AND BRUCE, N. Global burden of acute lower respiratory infections due to respiratory syncytial virus in young children: a systematic review and meta-analysis. *The Lancet* 375, 9725 (Apr. 2010), 1545–1555.
- [146] NEAL, P. Stochastic and Deterministic Analysis of SIS Household Epidemics. *Advances in applied probability* 38, 4 (2006), 943–968.
- [147] NOKES, D. J., MCLEAN, A. R., ANDERSON, R. M., AND GRABOWSKY, M. Measles Immunization Strategies for Countries with High Transmission

- Rates: Interim Guidelines Predicted Using a Mathematical Model. *International Journal of Epidemiology* 19, 3 (1990), 703–710.
- [148] NOKES, D. J., NGAMA, M., BETT, A., ABWAO, J., MUNYWOKI, P., ENGLISH, M., SCOTT, J. A., CANE, P. A., AND MEDLEY, G. F. Incidence and Severity of Respiratory Syncytial Virus Pneumonia in Rural Kenyan Children Identified through Hospital Surveillance. *Clin Infect Dis* (2009).
 - [149] NOKES, D. J., OKIRO, E. A., NGAMA, M., OCHOLA, R., WHITE, L. J., SCOTT, P. D., ENGLISH, M., CANE, P. A., AND MEDLEY, G. F. Respiratory syncytial virus infection and disease in infants and young children observed from birth in Kilifi District, Kenya. *Clin Infect Dis* 46, 1 (2008), 50–57.
 - [150] NOKES, D. J., OKIRO, E. A., NGAMA, M., WHITE, L. J., OCHOLA, R., SCOTT, P. D., CANE, P. A., AND MEDLEY, G. F. Respiratory syncytial virus epidemiology in a birth cohort from Kilifi district, Kenya: infection during the first year of life. *J Infect Dis* 190, 10 (2004), 1828–1832.
 - [151] NOKES, D. J., AND SWINTON, J. The control of childhood viral infections by pulse vaccination. *IMA journal of mathematics applied in medicine and biology* 12, 1 (Jan. 1995), 29–53.
 - [152] OCHOLA, R., SANDE, C., FEGAN, G., SCOTT, P. D., MEDLEY, G. F., CANE, P. A., AND NOKES, D. J. The Level and Duration of RSV-Specific Maternal IgG in Infants in Kilifi Kenya. *PLoS ONE* 4, 12 (2009), e8088.
 - [153] OCHOLA-OPIYO, B. R., AND NOKES, D. J. Passive and acquired immunity to Respiratory Syncytial Virus in young children in rural Kenya, 2007.
 - [154] OGILVIE, M. M., VATHENEN, A. S., RADFORD, M., CODD, J., AND KEY, S. Maternal antibody and respiratory syncytial virus infection in infancy. *J Med Virol* 7, 4 (1981), 263–271.
 - [155] OHUMA, E. O., OKIRO, E. A., OCHOLA, R., SANDE, C. J., CANE, P. A., MEDLEY, G. F., BOTTOMLEY, C., AND NOKES, D. J. The natural history

of respiratory syncytial virus in a birth cohort: the influence of age and previous infection on reinfection and disease. *American journal of epidemiology* 176, 9 (Nov. 2012), 794–802.

- [156] OKIRO, E. A. *Transmission Dynamics of Respiratory Syncytial Virus within the Household and in the Community*. PhD thesis, The Open University, 2007.
- [157] OKIRO, E. A., NGAMA, M., BETT, A., CANE, P. A., MEDLEY, G. F., AND JAMES NOKES, D. Factors associated with increased risk of progression to respiratory syncytial virus-associated pneumonia in young Kenyan children. *Trop Med Int Health* 13, 7 (2008), 914–926.
- [158] OKIRO, E. A., WHITE, L. J., NGAMA, M., CANE, P. A., MEDLEY, G. F., AND NOKES, D. J. Duration of shedding of respiratory syncytial virus in a community study of Kenyan children. *BMC Infect Dis* 10 (2010), 15.
- [159] OLSON, M., HARTWIG, S., AND VARGA, S. The number of respiratory syncytial virus (RSV)-specific memory CD8 T cells in the lung is critical for their ability to inhibit RSV vaccine-enhanced pulmonary eosino- philia. *The Journal of Immunology* 181, 11 (2008), 7958–7968.
- [160] PARROTT, R. H., KIM, H. W. H. A., ARROBIO, J. O., DAVID, S., MURPHY, B. R., BRANDT, C. D., CAMARGO, E. N. A., AND CHANOCK, R. M. Epidemiology of Respiratory Syncytial Virus infection in Washington, DC: II infection and disease with respect to age, immunologic status, race and sex. *American journal of epidemiology* 08, 4 (1973), 289–300.
- [161] PIEDRA, P. A. Future directions in vaccine prevention of respiratory syncytial virus. *Pediatr Infect Dis J* 21, 5 (2002), 482–487.
- [162] PIEDRA, P. A., JEWELL, A. M., CRON, S. G., ATMAR, R. L., AND PAUL GLEZEN, W. Correlates of immunity to respiratory syncytial virus (RSV) associated-hospitalization: establishment of minimum protective threshold levels of serum neutralizing antibodies. *Vaccine* 21, 24 (2003), 3479–3482.

- [163] PITZER, V., AND LIPSITCH, M. Exploring the relationship between incidence and the average age of infection during seasonal epidemics. *Journal of Theoretical Biology* 260, 2 (2009), 175–185.
- [164] PITZER, V. E., VIBOUD, C., SIMONSEN, L., STEINER, C., PANOZZO, C. A., ALONSO, W. J., MILLER, M. A., GLASS, R. I., GLASSER, J. W., PARASHAR, U. D., AND GRENFELL, B. T. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* 325, 5938 (2009), 290–294.
- [165] POLACK, F. P., TENG, M. N., L. COLLINS, P., PRINCE, G. A., EXNER, M., REGELE, H., LIRMAN, D. D., RABOLD, R., HOFFMAN, S. J., KARP, C. L., KLEEGERGER, S. R., WILLS-KARP, M., AND KARRON, R. A. A Role for Immune Complexes in Enhanced Respiratory Syncytial Virus Disease. *Journal of Experimental Medicine* 196, 6 (Sept. 2002), 859–865.
- [166] READ, J. M., EAMES, K. T., AND EDMUNDS, W. J. Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface* 5, 26 (2008), 1001–1007.
- [167] READ, J. M., EDMUNDS, W. J., RILEY, S., LESSLER, J., AND CUMMINGS, D. A. T. Close encounters of the infectious kind: methods to measure social mixing behaviour. *Epidemiology and infection* 140, 12 (Dec. 2012), 2117–30.
- [168] REESE, P. E., AND MARCHETTE, N. J. Respiratory syncytial virus infection and prevalence of subgroups A and B in Hawaii. *Journal of clinical microbiology* 29, 11 (Nov. 1991), 2614–5.
- [169] REICHERT, T. A., SUGAYA, N., FEDSON, D. S., GLEZEN, W. P., SIMONSEN, L., AND TASHIRO, M. The Japanese experience with vaccinating schoolchildren against influenza. *N Engl J Med* 344, 12 (2001), 889–896.
- [170] RESTIF, O., AND GRENFELL, B. T. Vaccination and the dynamics of immune evasion. *J R Soc Interface* 4, 12 (2007), 143–153.

- [171] ROBERTSON, S. E., ROCA, A., ALONSO, P., SIMOES, E. A. F., KARTASMITA, C. B., OLALEYE, D. O., ODAIBO, G. N., COLLINSON, M., VENTER, M., ZHU, Y., AND WRIGHT, P. F. Respiratory syncytial virus infection: denominator-based studies in Indonesia, Mozambique, Nigeria and South Africa. *Bulletin of the World Health Organization* 82, 12 (Dec. 2004), 914–22.
- [172] ROCA, A., ABACASSAMO, F., LOSCERTALES, M.-P., QUINTÓ, L., GÓMEZ-OLIVÉ, X., FENWICK, F., SAIZ, J. C., TOMS, G., AND ALONSO, P. L. Prevalence of respiratory syncytial virus IgG antibodies in infants living in a rural area of Mozambique. *Journal of medical virology* 67, 4 (Aug. 2002), 616–23.
- [173] ROHANI, P., ZHONG, X., AND KING, A. A. Contact Network Structure Explains the Changing Epidemiology of Pertussis. *Science* 982, 330 (2010).
- [174] ROSS, C. A. C., PINKERTON, I. W., AND ASSAAD, F. F. Pathogenesis of Respiratory Syncytial Virus Diseases in Infancy. *Archives of Disease in Childhood* 46 (1971), 702–704.
- [175] ROSS, J. V., HOUSE, T., AND KEELING, M. J. Calculation of Disease Dynamics in a Population of Households. *PLoS ONE* 5, 3 (2010), e9666.
- [176] RUE, H. V., MARTINO, S., AND CHOPIN, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 71, 2 (Apr. 2009), 319–392.
- [177] RUUSKANEN, O., LAHTI, E., JENNINGS, L. C., AND MURDOCH, D. R. Viral pneumonia. *Lancet* 6736, 10 (Mar. 2011), 1–12.
- [178] SANDE, C. J., MUTUNGA, M. N., MEDLEY, G. F., CANE, P. A., AND NOKES, D. J. Group- and genotype-specific neutralizing antibody responses against respiratory syncytial virus in infants and young children with severe pneumonia. *The Journal of infectious diseases* 207, 3 (Feb. 2013), 489–92.

- [179] SAUX, N. L., GABOURY, I., AND MACDONALD, N. Maternal respiratory syncytial virus antibody titers: season and children matter. *The Pediatric infectious ...* 22, 6 (2003), 563–564.
- [180] SCHENZLE, D. An Age-Structured Model of Pre- and Post-Vaccination Measles Transmission. *Mathematical Medicine and Biology* 1, 2 (1984), 169–191.
- [181] SCHICKLI, J. H., DUBOVSKY, F., AND TANG, R. S. Challenges in developing a pediatric RSV vaccine. *Human vaccines* 5, 9 (Sept. 2009), 582–91.
- [182] SCOTT, J. A. G., BAUNI, E., MOISI, J. C., OJAL, J., GATAKAA, H., NYUNDO, C., MOLYNEUX, C. S., KOMBE, F., TSOFA, B., MARSH, K., PESHU, N., AND WILLIAMS, T. N. Profile: The Kilifi Health and Demographic Surveillance System (KHDSS). *International journal of epidemiology* 41, 3 (June 2012), 650–7.
- [183] SCOTT, P., OCHOLA, R., AND SANDE, C. Comparison of Strain-Specific Antibody Responses During Primary and Secondary Infections With Respiratory Syncytial Virus. *Journal of Medical Virology* 1950, August (2007), 1943–1950.
- [184] SCOTT, P. D., OCHOLA, R., NGAMA, M., OKIRO, E. A., JAMES NOKES, D., MEDLEY, G. F., AND CANE, P. A. Molecular analysis of respiratory syncytial virus reinfections in infants from coastal Kenya. *J Infect Dis* 193, 1 (2006), 59–67.
- [185] SELWYN, B. J. The Epidemiology of Acute Respiratory Tract Infection in Young Children: Comparison of Findings from Several Developing Countries. *Reviews of Infectious Diseases* 12 (1990), S870–S888.
- [186] SHAMAN, J., AND KOHN, M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proceedings of the National Academy of Sciences of the United States of America* 106, 9 (Mar. 2009), 3243–8.
- [187] SHINOFF, J. J., O'BRIEN, K. L., THUMAR, B., SHAW, J. B., REID, R., HUA, W., SANTOSHAM, M., AND KARRON, R. A. Young infants

can develop protective levels of neutralizing antibody after infection with respiratory syncytial virus. *The Journal of infectious diseases* 198, 7 (Oct. 2008), 1007–15.

- [188] SIMOES, E. A. F. Environmental and demographic risk factors for respiratory syncytial virus lower respiratory tract disease. *The Journal of Pediatrics* 143, 5, Supplement 1 (2003), 118–126.
- [189] STENSALLE, L. G. An epidemiological study of respiratory syncytial virus associated hospitalizations in Denmark. *Respiratory research* 3 Suppl 1 (Jan. 2002), S34–9.
- [190] STENSALLE, L. G., DEVASUNDARAM, J. K., AND SIMOES, E. A. Respiratory syncytial virus epidemics: the ups and downs of a seasonal virus. *The Pediatric infectious disease journal* 22, 2 Suppl (Feb. 2003), S21–32.
- [191] STRANGERT, K. Respiratory illness in preschool children with different forms of day care., Feb. 1976.
- [192] SULLENDER, W. M. Respiratory syncytial virus genetic and antigenic diversity. *Clinical microbiology reviews* 13, 1 (Jan. 2000), 1–15, table of contents.
- [193] SUTMÖLLER, F., FERRO, Z. P., ASENSI, M. D., FERREIRA, V., MAZZEI, I. S., AND CUNHA, B. L. Etiology of acute respiratory tract infections among children in a combined community and hospital study in Rio de Janeiro. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 20, 4 (May 1995), 854–60.
- [194] TANG, R. S., SPAETE, R. R., THOMPSON, M. W., MACPHAIL, M., GUZZETTA, J. M., RYAN, P. C., REISINGER, K., CHANDLER, P., HILTY, M., WALKER, R. E., GOMEZ, M. M., AND LOSONSKY, G. A. Development of a PIV-vectored RSV vaccine: preclinical evaluation of safety, toxicity, and enhanced disease and initial clinical testing in healthy adults. *Vaccine* 26, 50 (2008), 6373–6382.

- [195] VAN DEN DRIESCHE, P., AND WATMOUGH, J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Mathematical Biosciences* 180, 1-2, 29–48.
- [196] WALLINGA, J., TEUNIS, P., AND KRETZSCHMAR, M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol* 164, 10 (2006), 936–944.
- [197] WALSH, E. E., PETERSON, D. R., AND FALSEY, A. R. Risk factors for severe respiratory syncytial virus infection in elderly persons. *The Journal of infectious diseases* 189, 2 (Jan. 2004), 233–8.
- [198] WARD, K. A., LAMBDEN, P. R., OGILVIE, M. M., AND WATT, P. J. Antibodies to respiratory syncytial virus polypeptides and their significance in human infection. *The Journal of general virology* 64 (Pt 9) (Sept. 1983), 1867–76.
- [199] WARIS, M., MAURICE, O. M., MUFSON, A., RUUSKANEN, O., AND HALONEN, P. Shedding of infectious virus and virus antigen during acute infection with respiratory syncytial virus. *Journal of Medical Virology* 38, 2 (1992), 111–116.
- [200] WARIS, M., AND WHITE, L. J. Seasonality of respiratory syncytial virus infection. *Clin Infect Dis* 43, 4 (2006), 541.
- [201] WEBER, A., WEBER, M., AND MILLIGAN, P. Modeling epidemics caused by respiratory syncytial virus (RSV). *Mathematical Biosciences* 172, 2 (2001), 95–113.
- [202] WEBER, M., MILLIGAN, P., HILTON, S., LAHAI, G., WHITTLE, H., MULHOLLAND, E. K., AND GREENWOOD, B. M. Risk factors for severe respiratory syncytial virus infection leading to hospital admission in children in the Western Region of The Gambia. *International Journal of Epidemiology* 28 (1999), 157–162.
- [203] WEBER, M. W., MILLIGAN, P., SANNEH, M., AWEMOYI, A., DAKOUR, R., SCHNEIDER, G., PALMER, A., JALLOW, M., OPARAOGU, A.,

- WHITTLE, H., MULHOLLAND, E. K., AND GREENWOOD, B. M. An epidemiological study of RSV infection in the Gambia. *Bulletin of the World Health Organization* 80, 7 (Jan. 2002), 562–8.
- [204] WEBER, M. W., MULHOLLAND, E. K., AND GREENWOOD, B. M. Respiratory syncytial virus infection in tropical and developing countries. *Tropical Medicine and International Health* 3 (1998), 268–280.
- [205] WHITE, L. J., MANDL, J. N., GOMES, M. G., BODLEY-TICKELL, A. T., CANE, P. A., PEREZ-BRENA, P., AGUILAR, J. C., SIQUEIRA, M. M., PORTES, S. A., STRALIOTTO, S. M., WARIS, M., NOKES, D. J., AND MEDLEY, G. F. Understanding the transmission dynamics of respiratory syncytial virus using multiple time series and nested models. *Math Biosci* 209, 1 (2007), 222–239.
- [206] WHITE, L. J., WARIS, M., CANE, P. A., NOKES, D. J., AND MEDLEY, G. F. The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection. *Epidemiol Infect* 133, 2 (2005), 279–289.
- [207] WILLIAMS, B. G., AND DYE, C. Maximum likelihood for parasitologists. *Parasitology Today* 10, 12 (1994), 489–493.
- [208] WRIGHT, P., KARRON, R., BELSHE, R., SHI, J., RANDOLPH, V. B., COLLINS, P. L., O'SHEA, A., GRUBER, W., AND MURPHY, B. R. The Absence of Enhanced Disease with Wild-Type Respiratory Syncytial Virus Infection Occurring After Receipt of Live, Attenuated, Respiratory Syncytial Virus Vaccines. *Vaccine* 25, 42 (2007), 7372–7378.
- [209] WU, J. T., RILEY, S., FRASER, C., AND LEUNG, G. M. Reducing the impact of the next influenza pandemic using household-based public health interventions. *PLoS Med* 3, 9 (2006), e361.

Appendix B

POLYMOD contacts for UK

Table B.1: Contact matrix of all reported contacts in Great Britain consisting of the average number of contact persons recorded per day per survey participant in the POLYMOD study

Age of contact	Age group of participant														
	0-4	5-9	10-14	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65-69	>70
0-4	1.92	0.65	0.41	0.24	0.46	0.73	0.67	0.83	0.24	0.22	0.36	0.2	0.2	0.26	0.13
5-9	0.95	6.64	1.09	0.73	0.61	0.75	0.95	1.39	0.9	0.16	0.3	0.22	0.5	0.48	0.2
10-14	0.48	1.31	6.85	1.52	0.27	0.31	0.48	0.76	1	0.69	0.32	0.44	0.27	0.41	0.33
15-19	0.33	0.34	1.03	6.71	1.58	0.73	0.42	0.56	0.85	1.16	0.7	0.3	0.2	0.48	0.63
20-24	0.45	0.3	0.22	0.93	2.59	1.49	0.75	0.63	0.77	0.87	0.88	0.61	0.53	0.37	0.33
25-29	0.79	0.66	0.44	0.74	1.29	1.83	0.97	0.71	0.74	0.85	0.88	0.87	0.67	0.74	0.33
30-34	0.97	1.07	0.62	0.5	0.88	1.19	1.67	0.89	1.02	0.91	0.92	0.61	0.76	0.63	0.27
35-39	1.02	0.98	1.26	1.09	0.76	0.95	1.53	1.5	1.32	1.09	0.83	0.69	1.02	0.96	0.2
40-44	0.55	1	1.14	0.94	0.73	0.88	0.82	1.23	1.35	1.27	0.89	0.67	0.94	0.81	0.8
45-49	0.29	0.54	0.57	0.77	0.97	0.93	0.57	0.8	1.32	1.87	0.61	0.8	0.61	0.59	0.57
50-54	0.33	0.38	0.4	0.41	0.44	0.85	0.6	0.61	0.71	0.95	0.74	1.06	0.59	0.56	0.57
55-59	0.31	0.21	0.25	0.33	0.39	0.53	0.68	0.53	0.55	0.51	0.82	1.17	0.85	0.85	0.33
60-64	0.26	0.25	0.19	0.24	0.19	0.34	0.4	0.39	0.47	0.55	0.41	0.78	0.65	0.85	0.57
65-69	0.09	0.11	0.12	0.2	0.19	0.22	0.13	0.3	0.23	0.13	0.21	0.28	0.36	0.7	0.6
> 70	0.14	0.15	0.21	0.1	0.24	0.17	0.15	0.41	0.5	0.71	0.53	0.76	0.47	0.74	1.47

Appendix C

Optimal sinks

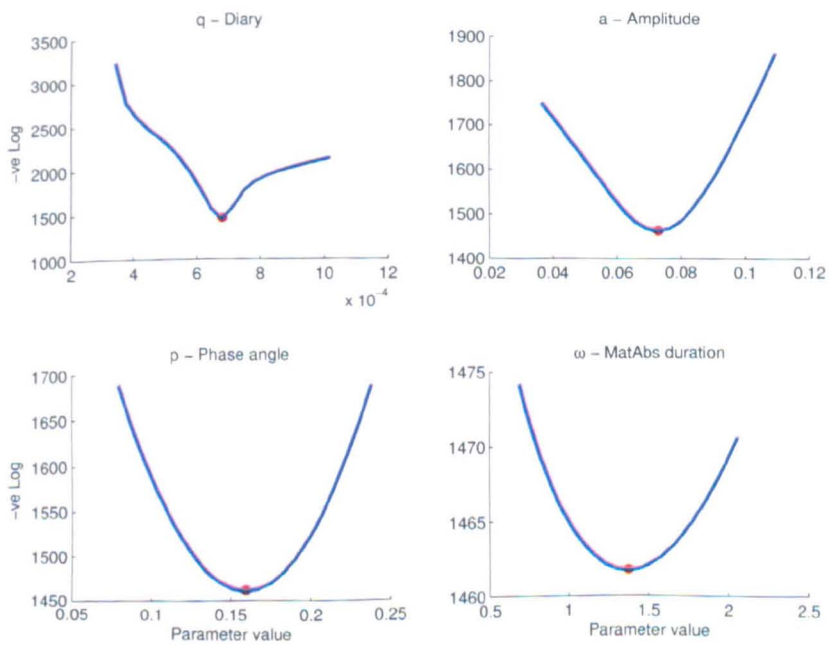


Figure C.1: Optimal sinks for the fitted parameters of the diary model. The red scatter plot represent the optimal value while the blue line shows the variation on either sides. To ensure optimality, all the scatter plots should fall within the lowest point in the graphs.

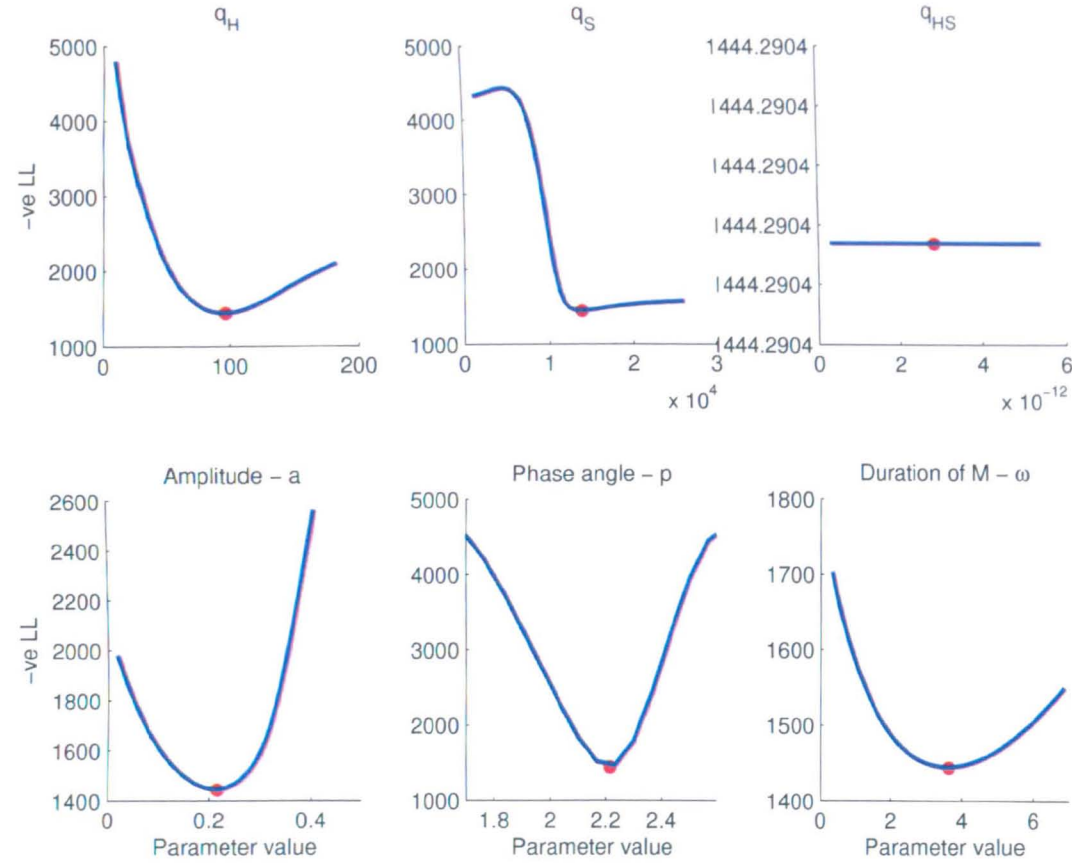


Figure C.2: Optimal sinks for the fitted parameters of the synthetic model. The red scatter plot represent the optimal value while the blue line shows the variation on either sides. To ensure optimality, all the scatter plots should fall within the lowest point in the graphs.

Appendix D

Parameter table

Table D.1 shows the parameters involved in the modelling exercise and how they were used. They are put under three headings 1) those that were part of the uncertainty and sensitivity analysis 2) those that were fitted during optimization and 3) those that were fixed during optimization. The shaded region shows the parameters that were part of the activity shown at the top of the column.

Table D.1: Shows the parameters and how they were used within the modelling work presented. All the fitted parameters were re-fitted for each of the Uncertainty and Sensitivity (U&S) scenario explored.

Symbol	Description	U&S	Fitted	Fixed
ω	Rate of decay of maternal antibodies			
$\beta_{i,j}$	Age-specific transmission parameter (WAIFW)			
σ_1, σ_2	Long-term immunity factor reducing the susceptibility of previously exposed individuals in S_1 and S_2			
ρ_0, ρ_1, ρ_2	Rate of waning of short-term immunity of recovered individuals, P_i $i=0,1,2$			
γ_0	Rate of recovery from primary infection, I_0			
γ_1, γ_2	Rate of recovery from secondary and tertiary infections, I_1, I_2			
$\alpha_0, \alpha_1, \alpha_2$	Factor reducing infectiousness of I_0, I_1 and I_2			
N_i	Initial population size where $i=1, \dots, 99$			
h_i	Proportion hospitalized where $i=1, \dots, 99$			
d	Risk of disease following infection where $i=1, \dots, 99$			
a	Relative amplitude			
ϕ	Phase angle - Peak of transmission			
q	Disease specific infectivity - Diary			
q_H, q_S, q_{HS}	Disease specific infectivity - Synthetic			
p	Number of maternal sub-classes			

Appendix E

Matlab code

Matlab code for the system of ODE. The function below returns the rate of change from the ODE system in Eqn 6.1 for use with the ode solver `ode45`.

```
1 function dy_dt = rate3DynamicS(time,y0,mu,birth,age,gamma,omega,...
    rho,eta,sigma,vs0,vs1,vs2,vsm,vsmm,beta,alpha,n,a,p,vaccine,...
    vage,nMclasses,nIclasses)
2
3 % The function solves the static age structured RSV model with ...
    the following as the input
4
5 % time: Time scalar passed by the ODE solver to the function
6 % y0: A vector of initial conditions updated by the solver at ...
    every time step
7 % mu: Vector of death rates
8 % birth: Vector of birth rates
9 % age: Rate of ageing, 1/(width of age class)
10 % gamma: Matrix containing the rate of recovery from I0,I1 and I2
11 % omega: Vector containing rate of loss of maternal antibody
12 % rho: Matrix containing the rate of waning of immunity from p0, ...
    p1 and p2 respectively
13 % eta: Homotopy parameters
14 % sigma: Matrix containing the parameters reducing the ...
    susceptibility of recovereds
15 % vs0: Vaccination coverage for S0 class into P0
16 % vs1: Vaccination coverage for S1 class into P1
```

```

17 % vs2: Vaccination coverage for S2 class into P2
18 % vsm: Vaccination coverage for M class at birth into the P0 ...
    class
19 % vsmm: Vaccination coverage for M class into the P0 class
20 % lambda: Static rate of infection
21 % n: Is the number of age classes required
22 % vage: Month at which to implement vaccination
23 % nMclasses: Number of desired maternal antibody sub-classes
24 % nIclasses: Number of desired primary infected sub-classes
25
26 % Function written and designed by:
27
28 % Author: Kinyanjui Timothy Muiruri
29 % Date: 2nd October 2009
30 % At: University of Warwick, UK, School of Biological Sciences
31 % email: timothykinyanjui@gmail.com
32
33 % Edited on the 2nd of Dec 2009
34 % Edited on the 5th of Jan 2010
35 % Edited on the 17th of Feb 2010 to include variable fertility ...
    rates.
36 % Edited on 28th April 2010 to include vaccination on S0, S1 and ...
    S2
37 % classes. Vaccination is assumed to elicit an immune response ...
    equivalent
38 % to a primary infection
39 % Edited on the 6th of May 2010 to include seasonality in the ...
    rate of
40 % infection
41 % Edited on 30th June 2010 to include time-independent rate of ...
    infection - foi
42 % Edited on 3rd March 2011 to include vaccination at the time of ...
    ageing
43 % rather than in the compartment itself
44 % Edited on the 19th of Sept 2011 to include variable M classes. ...
    This is
45 % to allow the decay of maternal antibody to follow a gamma ...
    distribution
46 % which is more realistic

```

```

47 % Edited on 8th Nov 2011 to include variable I0 su-classes. This ...
    allows
48 % the recovery to follow a gamma distribution which is more ...
    realistic
49
50 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
51
52 % Reshape y0 to be a matrix, with columns being equal to the ...
    number of
53 % epidemiological classes and rows being equal to the number of ...
    age
54 % classes
55 y0=reshape(y0,n,(nMclasses+1+nIclasses+7));
56
57 %% Find the sum in each age cohort
58 N=sum(y0,2);
59
60 % Implement vaccination at a desired time point
61 if time>20 % Change back to 20
62
63     % Vaccinate primary susceptibles
64     vs0(vage)=vaccine;
65
66     % Vaccinate secondary susceptibles
67     vs1(vage)=vaccine;
68
69     % Vaccinate tertiary susceptibles
70     vs2(vage)=vaccine;
71
72 end
73
74 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
75 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% Force of infection %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
76
77 % Calculate the indices to help in calculating the Force of ...
    Infection
78 % Density vs Frequency dependent? May be somewhere in between.
79 k1 = nMclasses+2;
80

```

```

81 k2 = nMclasses+1+nIclasses;
82
83 k3 = nMclasses+1+nIclasses+3;
84
85 k4 = nMclasses+1+nIclasses+6;
86
87 k5 = 1+(a*cos(2*pi*(time-p)));
88
89 % Frequency dependent
90 k6 = ((sum(y0(:,k1:k2),2).*alpha(:,1)) + (y0(:,k3).*alpha(:,2)) + ...
      (y0(:,k4).*alpha(:,3)))./N;
91
92 % Calculate the force of infection
93 lambda = k5*((beta')*k6);
94
95 %%%%%%%%%%%%%%% End of FOI calculation %%%%%%%%%%%%%%%
96 %%%%%%%%%%%%%%%
97
98
99 %%%%%%%%%%%%%%%
100 %% Write down the differential equations to solve the system %%
101
102 % Preallocate for code optimization
103 dy_dt=ones(n, (nMclasses+1+nIclasses+7));
104
105 % Pre-allocate force of infection
106 foi = lambda*ones(1,3);
107
108 % Set of equations for the first age class
109
110 % Generate the M-subclasses
111 % dy_dt(1,1) = (sum(birth.*N)*(1-vsm(1))-y0(1,1)*(mu(1)+age(1)+(...
      omega(1)*nMclasses));
112 dy_dt(1,1) = (birth*(1-vsm(1))-y0(1,1)*(mu(1)+age(1)+(omega(1)*...
      nMclasses));
113
114 % Subsequent M classes from the second to the last one
115 dy_dt(1,2:nMclasses) = (y0(1,1:(nMclasses-1))*omega(1)*nMclasses)-...
      y0(1,2:nMclasses)*(mu(1)+age(1)+(omega(1)*nMclasses));

```

```

116
117 % Primary susceptible class
118 dy_dt(1, (nMclasses+1)) = (omega(1) * nMclasses * y0(1, nMclasses)) - y0(1, (...
    nMclasses+1)) * (mu(1) + age(1) + foi(1, 1));
119
120 % Generate the primary infecteds sub-classes
121 dy_dt(1, (nMclasses+1+1)) = (foi(1, 1) * y0(1, (nMclasses+1))) - y0(1, (...
    nMclasses+1+1)) * (age(1) + (gamma(1, 1) * nIclasses) + mu(1));
122
123 % Generate subsequent I0 sub-classes to the last one
124 dy_dt(1, (nMclasses+1+2):(nMclasses+1+nIclasses)) = y0(1, (nMclasses...
    +1+1):(nMclasses+1+(nIclasses-1))) * (gamma(1, 1) * nIclasses) - y0...
    (1, (nMclasses+1+2):(nMclasses+1+nIclasses)) * (age(1) + mu(1) + (...
    gamma(1, 1) * nIclasses));
125
126 % Primary recovers class
127 % dy_dt(1, (nMclasses+1+nIclasses+1)) = (gamma(1, 1) * nIclasses * y0(1, (...
    nMclasses+1+nIclasses))) - y0(1, (nMclasses+1+nIclasses+1)) * (mu(1) ...
    + age(1) + rho(1, 1) + (eta(1, 1) * sigma(1, 1) * foi(1, 2))) + (sum(birth.*N)...
    * vsm(1));
128 dy_dt(1, (nMclasses+1+nIclasses+1)) = (gamma(1, 1) * nIclasses * y0(1, (...
    nMclasses+1+nIclasses))) - y0(1, (nMclasses+1+nIclasses+1)) * (mu(1) ...
    + age(1) + rho(1, 1) + (eta(1, 1) * sigma(1, 1) * foi(1, 2))) + (birth * vsm(1)) ...
    ;
129
130 % Secondary Susceptibles class
131 dy_dt(1, (nMclasses+1+nIclasses+2)) = (rho(1, 1) * y0(1, (nMclasses+1+...
    nIclasses+1))) - y0(1, (nMclasses+1+nIclasses+2)) * (mu(1) + age(1) + (...
    sigma(1, 1) * foi(1, 2)));
132
133 % Secondary infecteds class
134 dy_dt(1, (nMclasses+1+nIclasses+3)) = (sigma(1, 1) * foi(1, 2) * y0(1, (...
    nMclasses+1+nIclasses+2))) + (eta(1, 1) * sigma(1, 1) * foi(1, 2) * y0(1, (...
    nMclasses+1+nIclasses+1))) - y0(1, (nMclasses+1+nIclasses+3)) * (mu...
    (1) + age(1) + gamma(1, 2));
135
136 % Secondary recovers class
137 dy_dt(1, (nMclasses+1+nIclasses+4)) = (gamma(1, 2) * y0(1, (nMclasses+1+...
    nIclasses+3))) - y0(1, (nMclasses+1+nIclasses+4)) * (mu(1) + age(1) + ...

```



```

rho(1,2)+(eta(1,2)*sigma(1,2)*foi(1,3)));
138
139 % S2 class
140 dy_dt(1,(nMclasses+1+nIclasses+5))=(rho(1,2)*y0(1,(nMclasses+1+...
    nIclasses+4)))+(rho(1,3)*y0(1,(nMclasses+1+nIclasses+7)))-y0...
    (1,(nMclasses+1+nIclasses+5))*(mu(1)+age(1)+(sigma(1,2)*foi...
    (1,3)));
141
142 % I2 class
143 dy_dt(1,(nMclasses+1+nIclasses+6))=(eta(1,2)*sigma(1,2)*foi(1,3)*...
    y0(1,(nMclasses+1+nIclasses+4)))+(sigma(1,2)*foi(1,3)*y0(1,...
    nMclasses+1+nIclasses+5)))+(eta(1,3)*sigma(1,3)*foi(1,3)*y0(1,...
    nMclasses+1+nIclasses+7))-y0(1,(nMclasses+1+nIclasses+6))*(mu...
    (1)+age(1)+gamma(1,3));
144
145 % P2 class
146 dy_dt(1,(nMclasses+1+nIclasses+7))=(gamma(1,3)*y0(1,(nMclasses+1+...
    nIclasses+6)))-y0(1,(nMclasses+1+nIclasses+7))*(mu(1)+rho(1,3)+...
    age(1)+(eta(1,3)*sigma(1,3)*foi(1,3)));
147
148
149
150 %% The else part of the equations goes here but without the for ...
    loop
151
152 % Generate the M-subclasses
153 dy_dt(2:n,1)=((age(1:(n-1)).*y0(1:(n-1),1)).*(1-vsmm(1:(n-1))))-y0...
    (2:n,1).*(mu(2:n)+age(2:n)+(omega(2:n).*nMclasses)));
154
155 % Subsequent M classes from the second to the last one - This line...
    worked after two hours of tinkering
156 dy_dt(2:n,2:nMclasses)=(((age(1:n-1)+ones(1,length(2:nMclasses))))...
    *y0(1:n-1,2:nMclasses)).*((1-vsmm(1:n-1))*ones(1,length(2:...
    nMclasses))))+(y0(2:n,1:(nMclasses-1)).*((omega(2:n)*nMclasses)...
    *ones(1,length(2:nMclasses))))-y0(2:n,2:nMclasses).*((mu(2:n)+...
    age(2:n)+(omega(2:n).*nMclasses))*ones(1,length(2:nMclasses)));
157
158 % Primary susceptible class

```

```

159 dy_dt(2:n, (nMclasses+1)) = (omega(2:n).*y0(2:n, nMclasses).*...
    nMclasses)+(age(1:(n-1)).*y0(1:n-1, (nMclasses+1)).*(1-vs0(1:n...
    -1)))-y0(2:n, (nMclasses+1)).*(mu(2:n)+age(2:n)+foi(2:n, 1));
160
161 % Generate the primary infecteds sub-classes
162 dy_dt(2:n, (nMclasses+1+1)) = (foi(2:n, 1).*y0(2:n, (nMclasses+1)))+(...
    age(1:n-1).*y0(1:n-1, (nMclasses+1+1))-y0(2:n, (nMclasses+1+1))...
    *(age(2:n)+(gamma(2:n, 1)*nIclasses)+mu(2:n));
163
164 % Generate subsequent I0 sub-classes to the last one
165 dy_dt(2:n, (nMclasses+1+2):(nMclasses+1+nIclasses)) = (age(1:n-1)*...
    ones(1, length((nMclasses+1+2):(nMclasses+1+nIclasses)))).*y0(1:...
    n-1, (nMclasses+1+2):(nMclasses+1+nIclasses))+y0(2:n, (nMclasses...
    +1+1):(nMclasses+1+(nIclasses-1))).*(gamma(2:n, 1)*nIclasses)*...
    ones(1, length((nMclasses+1+1):(nMclasses+1+(nIclasses-1)))))-y0...
    (2:n, (nMclasses+1+2):(nMclasses+1+nIclasses)).*((age(2:n)+mu(2:...
    n)+(gamma(2:n, 1)*nIclasses))*ones(1, length((nMclasses+1+2):(...
    nMclasses+1+nIclasses)))));
166
167 % Primary recovers class
168 dy_dt(2:n, (nMclasses+1+nIclasses+1))=(gamma(2:n, 1).*nIclasses.*y0...
    (2:n, (nMclasses+1+nIclasses)))+(age(1:n-1).*y0(1:n-1, (nMclasses...
    +1+nIclasses+1))-y0(2:n, (nMclasses+1+nIclasses+1)).*(mu(2:n)+...
    age(2:n)+rho(2:n, 1)+(eta(2:n, 1).*sigma(2:n, 1).*foi(2:n, 2)))+( ...
    age(1:n-1).*y0(1:n-1, nMclasses+1+nIclasses)).*vsmm(1:n-1))+(age...
    (1:n-1).*y0(1:n-1, (nMclasses+1)).*vs0(1:n-1));
169
170 % Secondary Susceptibles class
171 dy_dt(2:n, (nMclasses+1+nIclasses+2))=(rho(2:n, 1).*y0(2:n, (...
    nMclasses+1+nIclasses+1)))+(age(1:n-1).*y0(1:n-1, (nMclasses+1+...
    nIclasses+2)).*(1-vs1(1:n-1))-y0(2:n, (nMclasses+1+nIclasses+2)...
    ).*(mu(2:n)+age(2:n)+(sigma(2:n, 1).*foi(2:n, 2))));
172
173 % Secondary infecteds class
174 dy_dt(2:n, (nMclasses+1+nIclasses+3))=(sigma(2:n, 1).*foi(2:n, 2).*y0...
    (2:n, (nMclasses+1+nIclasses+2)))+(age(1:n-1).*y0(1:n-1, (...
    nMclasses+1+nIclasses+3)))+(eta(2:n, 1).*sigma(2:n, 1).*foi(2:n...
    , 2).*y0(2:n, (nMclasses+1+nIclasses+1))-y0(2:n, (nMclasses+1+...
    nIclasses+3)).*(mu(2:n)+age(2:n)+gamma(2:n, 2)));

```

```

175
176 % Secondary recovers class
177 dy_dt(2:n, (nMclasses+1+nIclasses+4))=(gamma(2:n,2).*y0(2:n, (...
    nMclasses+1+nIclasses+3)))+(age(1:n-1).*y0(1:n-1, (nMclasses+1+...
    nIclasses+4)))-y0(2:n, (nMclasses+1+nIclasses+4)).*(mu(2:n)+age...
    (2:n)+rho(2:n,2)+(eta(2:n,2).*sigma(2:n,2).*foi(2:n,3)))+(age...
    (1:n-1).*y0(1:n-1, (nMclasses+1+nIclasses+2)).*vs1(1:n-1));
178
179 % S2 class
180 dy_dt(2:n, (nMclasses+1+nIclasses+5))=(rho(2:n,2).*y0(2:n, (...
    nMclasses+1+nIclasses+4)))+(rho(2:n,3).*y0(2:n, (nMclasses+1+...
    nIclasses+7)))+(age(1:n-1).*y0(1:n-1, (nMclasses+1+nIclasses+5))...
    .*(1-vs2(1:n-1)))-y0(2:n, (nMclasses+1+nIclasses+5)).*(mu(2:n)+...
    age(2:n)+(sigma(2:n,2).*foi(2:n,3)));
181
182 % I2 class
183 dy_dt(2:n, (nMclasses+1+nIclasses+6))=(eta(2:n,2).*sigma(2:n,2).*...
    foi(2:n,3).*y0(2:n, (nMclasses+1+nIclasses+4)))+(age(1:n-1).*y0...
    (1:n-1, (nMclasses+1+nIclasses+6)))+(sigma(2:n,2).*foi(2:n,3).*...
    y0(2:n, (nMclasses+1+nIclasses+5)))+(eta(2:n,3).*sigma(2:n,3).*...
    foi(2:n,3).*y0(2:n, (nMclasses+1+nIclasses+7)))-y0(2:n, (...
    nMclasses+1+nIclasses+6)).*(mu(2:n)+age(2:n)+gamma(2:n,3));
184
185 % P2 class
186 dy_dt(2:n, (nMclasses+1+nIclasses+7))=(gamma(2:n,3).*y0(2:n, (...
    nMclasses+1+nIclasses+6)))+(age(1:n-1).*y0(1:n-1, (nMclasses+1+...
    nIclasses+7)))-y0(2:n, (nMclasses+1+nIclasses+7)).*(mu(2:n)+rho...
    (2:n,3)+age(2:n)+(eta(2:n,3).*sigma(2:n,3).*foi(2:n,3)))+(age...
    (1:n-1).*y0(1:n-1, (nMclasses+1+nIclasses+5)).*vs2(1:n-1));
187
188
189 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%% End of differential equations %%%%%%%%%
190 %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
191
192
193
194 %% Return a column vector for use with the ODE solver
195 dy_dt=reshape(dy_dt, (n*(nMclasses+1+nIclasses+7)),1);
196

```

197 return